



(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)



1-0052103

(51)^{2020.01} H04N 19/70

(13) B

(21) 1-2022-02462

(22) 24/09/2020

(86) PCT/CN2020/117557 24/09/2020

(87) WO 2021/057869 01/04/2021

(30) PCT/CN2019/107594 24/09/2019 CN

(45) 25/09/2025 450

(43) 25/07/2022 412A

(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)

Huawei Administration Building, Bantian, Longgang District, Shenzhen, Guangdong
518129, P. R. China

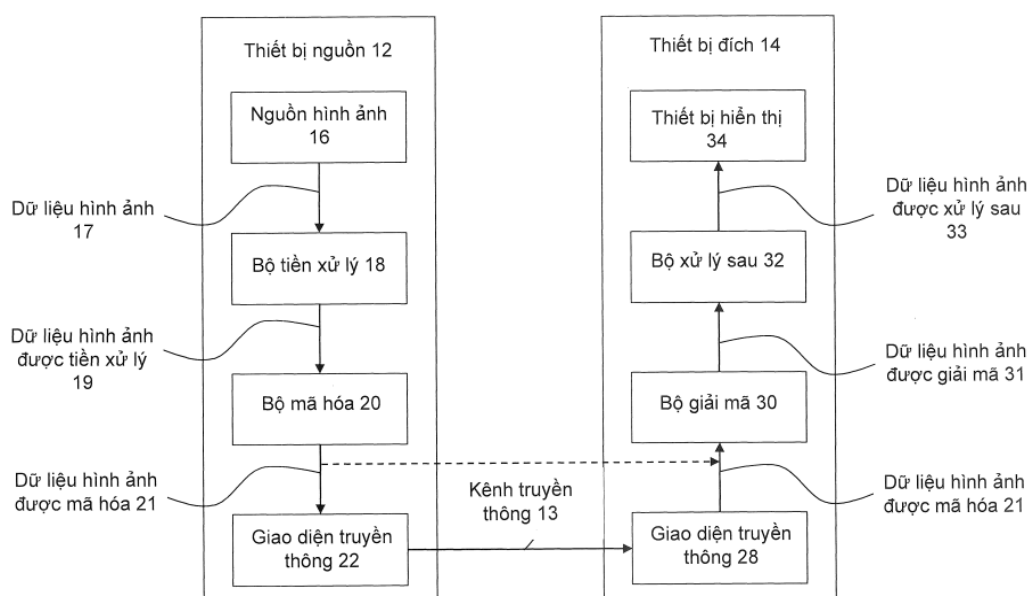
(72) MA, Xiang (CN); YANG, Haitao (CN).

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

(54) PHƯƠNG PHÁP GIẢI MÃ DÒNG BIT VIDEO ĐƯỢC LẬP MÃ, PHƯƠNG PHÁP
MÃ HÓA DÒNG BIT VIDEO ĐƯỢC LẬP MÃ, THIẾT BỊ GIẢI MÃ, THIẾT BỊ
MÃ HÓA, PHƯƠNG TIỆN BẮT BIẾN CÓ THỂ ĐỌC ĐƯỢC BỞI MÁY TÍNH

(21) 1-2022-02462

(57) Sáng chế đề cập đến phương pháp giải mã dòng bit video được lập mã, phương pháp mã hóa dòng bit video được lập mã, thiết bị giải mã, thiết bị mã hóa, phương tiện bất biến có thể đọc được bằng máy tính, phương tiện lưu trữ bất biến, phương pháp giải mã dòng bit video được lập mã bao gồm: thu thành phần cú pháp mức tham số chuỗi (SPS - sequence parameter set) từ dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS - video parameter set) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; thu, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP, inter-layer reference picture) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không; và dự báo một hoặc nhiều hình ảnh được lập mã dựa trên giá trị của thành phần cú pháp cho phép liên lớp.



Lĩnh vực kỹ thuật được đề cập

Sáng chế (các phương án của sáng chế) nói chung đề cập đến lĩnh vực xử lý hình ảnh và cụ thể hơn là dự báo liên ảnh.

Tình trạng kỹ thuật của sáng chế

Việc lập mã video (mã hóa và giải mã video) được sử dụng trong một loạt các ứng dụng video kỹ thuật số, ví dụ tivi kỹ thuật số phát quảng bá, truyền video qua các mạng internet và di động, các ứng dụng đàm thoại theo thời gian thực chẳng hạn như trò chuyện qua video, hội thảo qua video, các đĩa DVD và Blu-ray, các hệ thống thu và biên tập nội dung video, và các máy quay xách tay có các ứng dụng bảo mật.

Lượng dữ liệu video cần để mô tả ngay cả một video tương đối ngắn có thể là đáng kể, điều này có thể gây ra các khó khăn khi dữ liệu sẽ được truyền trực tuyến hoặc nếu không thì được truyền thông qua mạng truyền thông có dung lượng băng thông giới hạn. Do đó, dữ liệu video thông thường được nén trước khi được truyền thông qua các mạng viễn thông hiện đại. Kích cỡ video cũng có thể là vấn đề khi video này được lưu trữ trên thiết bị lưu trữ do các tài nguyên bộ nhớ có thể bị giới hạn. Các thiết bị nén video thường sử dụng phần mềm và/hoặc phần cứng ở nguồn để mã hóa dữ liệu video trước khi truyền hoặc lưu trữ, nhờ đó giảm lượng dữ liệu cần để thể hiện các ảnh video kỹ thuật số. Dữ liệu được nén sau đó được thu nhận ở đích bởi thiết bị giải nén video mà giải mã dữ liệu video. Với các tài nguyên mạng hạn chế và nhu cầu về chất lượng video cao hơn ngày một gia tăng, nên mong muốn có các kỹ thuật nén và giải nén cải tiến mà cải thiện tỷ lệ nén mà ít hoặc không hy sinh chất lượng hình ảnh.

Bản chất kỹ thuật của sáng chế

Các phương án của sáng chế đề xuất các thiết bị và các phương pháp mã hóa và giải mã theo các điểm yêu cầu bảo hộ độc lập.

Các mục đích ở trên và các mục đích khác đạt được bằng các đối tượng của các điểm yêu cầu bảo hộ độc lập. Ngoài ra các cách thức thực hiện khác là rõ ràng từ các điểm yêu cầu bảo hộ độc lập, phần mô tả và các hình vẽ.

Các phương án cụ thể được phác họa trong các điểm yêu cầu bảo hộ độc lập đi kèm, với các phương án khác có trong các điểm yêu cầu bảo hộ phụ thuộc.

Theo khía cạnh thứ nhất sáng chế đề xuất phương pháp giải mã dòng bit video được lập mã. Phương pháp này được thực hiện bởi thiết bị giải mã. Phương pháp này bao gồm: thu thành phần cú pháp mức tham số chuỗi (SPS - sequence parameter set) từ dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS - video parameter set) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; thu, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP, inter-layer reference picture) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không; và dự báo một hoặc nhiều hình ảnh được lập mã dựa trên giá trị của thành phần cú pháp cho phép liên lớp.

Khi dự báo liên ảnh, hình ảnh được lập mã và hình ảnh tham chiếu của hình ảnh được lập mã thuộc về các lớp khác nhau, trong đó các lớp khác nhau này có thể tương ứng với độ phân giải khác nhau, và độ phân giải không gian thấp có thể được sử dụng làm tham chiếu cho độ phân giải không gian cao. Thành phần cú pháp cho phép liên lớp xác định liệu dự báo liên ảnh có được cho phép hay không, do đó các thành phần cú pháp liên quan đến dự báo liên ảnh không cần phải được báo hiệu khi thành phần cú pháp cho phép liên lớp xác định dự báo liên ảnh bị ngắt, do đó có thể giảm tốc độ bit. Ngoài ra, tập tham số video (VPS) được dùng cho nhiều lớp, không có VPS nào được tham chiếu đến bởi SPS có nghĩa là không cần nhiều lớp khi giải mã các hình ảnh liên quan đến SPS, nói cách khác, chỉ có một lớp sẽ được sử dụng khi giải mã các hình ảnh liên quan đến SPS. Dự báo liên ảnh không thể được thực hiện khi chỉ có một lớp, do đó không báo hiệu thành phần cú pháp cho phép liên lớp khi không có VPS nào được tham chiếu đến bởi SPS sẽ giảm thêm tốc độ bit.

Trong đó dòng bit là chuỗi các bit tạo ra một hoặc nhiều chuỗi video được mã hóa (CVS - coded video sequence).

Trong đó chuỗi video được mã hóa (CVS) là chuỗi các AU (đơn vị truy nhập - access unit).

Trong đó chuỗi video lớp được mã hóa (CLVS - coded layer video sequence) là chuỗi các PU (đơn vị hình ảnh - picture unit) có cùng giá trị `nuh_layer_id`.

Trong đó đơn vị truy nhập (AU) là tập các PU mà thuộc về các lớp khác nhau và chứa các hình ảnh được lập mã liên kết có cùng một thời gian để xuất ra từ DPB (decoded picture buffer – bộ đệm hình ảnh được giải mã).

Trong đó đơn vị hình ảnh (PU) là tập các đơn vị NAL (network abstraction layer – lớp trừu tượng mạng) mà liên kết với nhau theo quy tắc phân loại cụ thể, liên tiếp theo thứ tự giải mã, và chứa chính xác một hình ảnh được lập mã.

Trong đó hình ảnh tham chiếu liên lớp (ILRP) là hình ảnh trong cùng một AU với hình ảnh hiện thời, với `nuh_layer_id` nhỏ hơn `nuh_layer_id` của hình ảnh hiện thời.

Trong đó SPS là cấu trúc cú pháp chứa các thành phần cú pháp mà áp dụng cho 0 hoặc nhiều hơn toàn bộ các CLVS.

Như vậy trong các cách thức thực hiện khả thi của phương pháp theo khía cạnh thứ nhất, trong đó VPS bao gồm các thành phần cú pháp mô tả thông tin dự báo liên lớp của các lớp trong chuỗi video được mã hóa (CVS), và SPS bao gồm thành phần cú pháp mức SPS và thành phần cú pháp cho phép liên lớp, trong đó CVS bao gồm một hoặc nhiều ILRP và một hoặc nhiều hình ảnh được lập mã.

Trong đó khi VPS được tham chiếu đến bởi SPS, VPS này bao gồm các thành phần cú pháp mô tả thông tin dự báo liên lớp của các lớp mà một hoặc nhiều ILRP và một hoặc nhiều hình ảnh được lập mã thuộc về.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ nhất hoặc khía cạnh thứ nhất như vậy, trong đó việc dự báo một hoặc nhiều hình ảnh được lập mã dựa trên giá trị của thành phần cú pháp cho phép liên lớp bao gồm: khi giá trị của thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã, dự báo một hoặc nhiều hình ảnh được lập mã bằng cách tham chiếu một hoặc nhiều ILRP, trong đó một hoặc nhiều ILRP này thu được dựa trên thông tin dự báo liên lớp được bao gồm trong VPS được tham chiếu đến bởi SPS.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ nhất hoặc khía cạnh thứ nhất như vậy, trong đó hình ảnh được lập mã và ILRP của hình ảnh được lập mã thuộc về các lớp khác nhau.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ nhất hoặc khía cạnh thứ nhất như vậy, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước còn xác định rằng

chuỗi video được mã hóa (CVS) bao gồm các hình ảnh được lập mã của chỉ một lớp.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ nhất hoặc khía cạnh thứ nhất như vậy, trong đó giá trị thiết đặt trước bằng 0.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ nhất hoặc khía cạnh thứ nhất như vậy, trong đó dự báo một hoặc nhiều hình ảnh được lập mã dựa trên giá trị của thành phần cú pháp cho phép liên lớp bao gồm: dự báo một hoặc nhiều hình ảnh được lập mã mà không cần tham chiếu bất kỳ ILRP nào trong trường hợp giá trị của thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều ILRP không được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã.

Theo khía cạnh thứ hai sáng chế đề xuất phương pháp mã hóa dòng bit video được lập mã. Phương pháp này được thực hiện bởi thiết bị mã hóa. Phương pháp này bao gồm: mã hóa thành phần cú pháp mức tập tham số chuỗi (SPS) trong dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; mã hóa, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp trong dòng bit, trong đó thành phần cú pháp cho phép liên lớp này xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không.

Trong các cách thức thực hiện khả thi của phương pháp theo khía cạnh thứ hai như vậy, trong đó VPS bao gồm các thành phần cú pháp mô tả thông tin dự báo liên lớp của các lớp trong chuỗi video được mã hóa (CVS), và SPS bao gồm thành phần cú pháp mức SPS và thành phần cú pháp cho phép liên lớp, trong đó CVS bao gồm một hoặc nhiều ILRP và một hoặc nhiều hình ảnh được lập mã.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ hai hoặc khía cạnh thứ hai như vậy, trong đó hình ảnh được lập mã và ILRP của hình ảnh được lập mã thuộc về các lớp khác nhau.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ hai hoặc khía cạnh thứ hai như vậy, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước còn xác định rằng

chuỗi video được mã hóa (CVS) bao gồm các hình ảnh được lập mã của chỉ một lớp.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ hai hoặc khía cạnh thứ hai như vậy, trong đó giá trị thiết đặt trước bằng 0.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ hai hoặc khía cạnh thứ hai như vậy, trong đó việc mã hóa thành phần cú pháp cho phép liên lớp trong dòng bit bao gồm: mã hóa thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều ILRP được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong dòng bit dựa trên việc xác định rằng một hoặc nhiều ILRP được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ hai hoặc khía cạnh thứ hai như vậy, trong đó việc mã hóa thành phần cú pháp cho phép liên lớp trong dòng bit bao gồm: mã hóa thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều ILRP không được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong dòng bit dựa trên việc xác định rằng một hoặc nhiều ILRP không được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã.

Theo khía cạnh thứ ba sáng chế đề xuất bộ giải mã dùng để giải mã dòng bit video được lập mã, bộ giải mã này bao gồm: bộ thu được tạo cấu hình để thu thành phần cú pháp mức tập tham số chuỗi (SPS) từ dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; bộ thu này còn được tạo cấu hình để thu, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không; và bộ dự báo được tạo cấu hình để dự báo một hoặc nhiều hình ảnh được lập mã dựa trên giá trị của thành phần cú pháp cho phép liên lớp.

Trong các cách thức thực hiện khả thi của phương pháp theo khía cạnh thứ ba như vậy, trong đó VPS bao gồm các thành phần cú pháp mô tả thông tin dự báo liên lớp của các lớp trong chuỗi video được mã hóa (CVS), và SPS bao gồm thành

phần cú pháp mức SPS và thành phần cú pháp cho phép liên lớp, trong đó CVS bao gồm một hoặc nhiều ILRP và một hoặc nhiều hình ảnh được lập mã.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ ba hoặc khía cạnh thứ ba như vậy, bộ dự báo được tạo cấu hình để dự báo, dưới dạng giá trị của thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã, một hoặc nhiều hình ảnh được lập mã bằng cách tham chiếu một hoặc nhiều ILRP, trong đó một hoặc nhiều ILRP thu được dựa trên thông tin dự báo liên lớp được bao gồm trong VPS được tham chiếu đến bởi SPS.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ ba hoặc khía cạnh thứ ba như vậy, trong đó hình ảnh được lập mã và ILRP của hình ảnh được lập mã thuộc về các lớp khác nhau.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ ba hoặc khía cạnh thứ ba như vậy, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước còn xác định rằng chuỗi video được mã hóa (CVS) bao gồm các hình ảnh được lập mã của chỉ một lớp.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ ba hoặc khía cạnh thứ ba như vậy, trong đó giá trị thiết đặt trước bằng 0.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ ba hoặc khía cạnh thứ ba như vậy, bộ dự báo được tạo cấu hình để dự báo một hoặc nhiều hình ảnh được lập mã mà không cần tham chiếu bất kỳ ILRP nào trong trường hợp giá trị của thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều ILRP không được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã.

Theo khía cạnh thứ tư sáng chế đề xuất bộ mã hóa dùng để mã hóa dòng bit video được lập mã, bộ mã hóa này bao gồm: bộ mã hóa thứ nhất được tạo cấu hình để mã hóa thành phần cú pháp mức tập tham số chuỗi (SPS) trong dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; bộ mã hóa thứ hai được tạo cấu hình để mã hóa, dưới dạng thành phần cú

pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp trong dòng bit, trong đó thành phần cú pháp cho phép liên lớp này xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không.

Trong các cách thức thực hiện khả thi của phương pháp theo khía cạnh thứ tư như vậy, bộ mã hóa còn bao gồm bộ xác định được tạo cấu hình để xác định liệu thành phần cú pháp mức SPS có lớn hơn giá trị thiết đặt trước hay không.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ tư hoặc khía cạnh thứ tư như vậy, trong đó VPS bao gồm các thành phần cú pháp mô tả thông tin dự báo liên lớp của các lớp trong chuỗi video được mã hóa (CVS), và SPS bao gồm thành phần cú pháp mức SPS và thành phần cú pháp cho phép liên lớp, trong đó CVS bao gồm một hoặc nhiều ILRP và một hoặc nhiều hình ảnh được lập mã.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ tư hoặc khía cạnh thứ tư như vậy, trong đó hình ảnh được lập mã và ILRP của hình ảnh được lập mã thuộc về các lớp khác nhau.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ tư hoặc khía cạnh thứ tư như vậy, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước còn xác định rằng chuỗi video được mã hóa (CVS) bao gồm các hình ảnh được lập mã của chỉ một lớp.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ tư hoặc khía cạnh thứ tư như vậy, trong đó giá trị thiết đặt trước bằng 0.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ tư hoặc khía cạnh thứ tư như vậy, trong đó bộ mã hóa thứ hai được tạo cấu hình để mã hóa thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều ILRP được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong dòng bit dựa trên việc xác định rằng một hoặc nhiều ILRP được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ tư hoặc khía cạnh thứ tư như vậy, trong đó

bộ mã hóa thứ hai được tạo cấu hình để mã hóa thành phần cú pháp cho phép liên lớp xác định một hoặc nhiều ILRP không được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong dòng bit dựa trên việc xác định rằng một hoặc nhiều ILRP không được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã.

Trong các cách thức thực hiện khả thi của phương pháp theo bất kỳ cách thức thực hiện nào trước đó của khía cạnh thứ tư hoặc khía cạnh thứ tư như vậy, bộ mã hóa còn bao gồm bộ xác định được tạo cấu hình để xác định liệu một hoặc nhiều ILRP có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không.

Phương pháp theo khía cạnh thứ nhất của sáng chế có thể được thực hiện bởi thiết bị theo khía cạnh thứ ba của sáng chế. Ngoài ra các đặc điểm và các cách thức thực hiện của phương pháp theo khía cạnh thứ ba của sáng chế tương ứng với các đặc điểm và cách thức thực hiện của thiết bị theo khía cạnh thứ nhất của sáng chế.

Phương pháp theo khía cạnh thứ hai của sáng chế có thể được thực hiện bởi thiết bị theo khía cạnh thứ tư của sáng chế. Ngoài ra các đặc điểm và các cách thức thực hiện của phương pháp theo khía cạnh thứ tư của sáng chế tương ứng với các đặc điểm và cách thức thực hiện của thiết bị theo khía cạnh thứ hai của sáng chế.

Phương pháp theo khía cạnh thứ hai có thể được mở rộng trong các cách thức thực hiện tương ứng với các cách thức thực hiện của thiết bị thứ nhất theo khía cạnh thứ nhất. Do đó, cách thức thực hiện của phương pháp này bao gồm (các) đặc điểm của của cách thức thực hiện tương ứng của thiết bị thứ nhất.

Các ưu điểm của các phương pháp theo khía cạnh thứ hai giống với các ưu điểm của các cách thức thực hiện tương ứng của thiết bị thứ nhất theo khía cạnh thứ nhất.

Theo khía cạnh thứ năm sáng chế đề xuất thiết bị giải mã dòng video bao gồm bộ xử lý và bộ nhớ. Bộ nhớ này lưu trữ các lệnh mà khiến bộ xử lý thực hiện phương pháp theo khía cạnh thứ nhất.

Theo khía cạnh thứ sáu sáng chế đề xuất thiết bị mã hóa dòng video bao gồm bộ xử lý và bộ nhớ. Bộ nhớ này lưu trữ các lệnh mà khiến bộ xử lý thực hiện phương pháp theo khía cạnh thứ hai.

Theo khía cạnh thứ bảy, sáng chế đề xuất phương tiện lưu trữ có thể đọc được bằng máy tính lưu trữ trên đó các lệnh mà khi được thực hiện sẽ khiến một hoặc nhiều bộ xử lý được tạo cấu hình để mã hóa dữ liệu video. Các lệnh này khiến

một hoặc nhiều bộ xử lý thực hiện phương pháp theo khía cạnh thứ nhất hoặc thứ hai hoặc bất kỳ phương án khả thi nào của khía cạnh thứ nhất hoặc thứ hai.

Theo khía cạnh thứ tám, sáng chế đề xuất chương trình máy tính bao gồm mã chương trình dùng để thực hiện phương pháp theo khía cạnh thứ nhất hoặc thứ hai hoặc bất kỳ phương án khả thi nào của khía cạnh thứ nhất hoặc thứ hai khi được thực hiện trên máy tính.

Theo khía cạnh thứ chín, sáng chế đề xuất phương tiện lưu trữ bất biến mà bao gồm dòng bit được lập mã được giải mã bởi thiết bị giải mã hình ảnh, dòng bit này được tạo ra bằng cách chia khung tín hiệu video hoặc tín hiệu ảnh thành nhiều khối, và bao gồm nhiều thành phần cú pháp, trong đó các thành phần cú pháp này bao gồm thành phần cú pháp cho phép liên lớp xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không với điều kiện là thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS.

Chi tiết về một hoặc nhiều phương án được trình bày trong các hình vẽ kèm theo và phần mô tả dưới đây. Các đặc điểm, các mục đích, và các ưu điểm sẽ là rõ ràng từ phần mô tả, các hình vẽ, và bộ yêu cầu bảo hộ.

Ngoài ra, sáng chế đề xuất các phương án sau.

Theo một phương án, sáng chế đề xuất phương pháp giải mã dòng bit video được lập mã, phương pháp này bao gồm:

phân tích cú pháp thành phần cú pháp thứ nhất xác định liệu lớp có chỉ số i có sử dụng dự báo liên lớp hay không, i là số nguyên và i lớn hơn 0;

khi điều kiện thứ nhất được thỏa mãn, phân tích cú pháp thành phần cú pháp thứ hai xác định liệu lớp có chỉ số j có phải là lớp tham chiếu trực tiếp đối với lớp có chỉ số i hay không, trong đó j là số nguyên, và j nhỏ hơn i và lớn hơn hoặc bằng 0, trong đó điều kiện thứ nhất bao gồm thành phần cú pháp thứ nhất xác định rằng lớp có chỉ số i có thể sử dụng dự báo liên lớp và i lớn hơn giá trị thiết đặt trước (ví dụ, 1);

dự báo hình ảnh của lớp có chỉ số i dựa trên giá trị của thành phần cú pháp thứ hai.

Theo một phương án, trong đó phương pháp này còn bao gồm:

khi điều kiện thứ hai được thỏa mãn, thì dự báo hình ảnh của lớp có chỉ số i bằng cách sử dụng lớp có chỉ số j làm lớp tham chiếu trực tiếp đối với lớp có chỉ số i , trong đó j là số nguyên, và j nhỏ hơn i và lớn hơn hoặc bằng 0, trong đó điều kiện thứ hai bao gồm thành phần cú pháp xác định rằng lớp có chỉ số i có thể sử dụng dự báo liên lớp và i bằng với giá trị thiết đặt trước.

Theo một phương án, trong đó phương pháp này còn bao gồm:

khi điều kiện thứ hai được thỏa mãn, thì việc xác định giá trị của thành phần cú pháp thứ hai xác định lớp có chỉ số j là lớp tham chiếu trực tiếp đối với lớp có chỉ số i .

Theo một phương án, trong đó hình ảnh của lớp có chỉ số i bao gồm hình ảnh lớp có chỉ số i hoặc hình ảnh liên quan đến lớp có chỉ số i .

Theo một phương án, sáng chế đề xuất phương pháp giải mã dòng bit video được lập mã, phương pháp này bao gồm:

phân tích cú pháp thành phần cú pháp xác định liệu lớp có chỉ số i có sử dụng dự báo liên lớp hay không, i là số nguyên và i lớn hơn 0;

khi điều kiện được thỏa mãn, thì dự báo hình ảnh của lớp có chỉ số i bằng cách sử dụng lớp có chỉ số j làm lớp tham chiếu trực tiếp đối với lớp có chỉ số i , trong đó j là số nguyên, và j bằng $i-1$, trong đó điều kiện này bao gồm thành phần cú pháp xác định rằng lớp có chỉ số i có thể sử dụng dự báo liên lớp.

Theo một phương án, trong đó hình ảnh của lớp có chỉ số i bao gồm hình ảnh lớp có chỉ số i hoặc hình ảnh liên quan đến lớp có chỉ số i .

Theo một phương án, sáng chế đề xuất phương pháp giải mã dòng bit video được lập mã, phương pháp này bao gồm:

phân tích cú pháp thành phần cú pháp xác định liệu ít nhất một hình ảnh tham chiếu dài hạn (LTRP, long-term reference picture) có được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong chuỗi video được mã hóa (CVS) hay không, trong đó mỗi hình ảnh của ít nhất một LTRP được chú thích là "được sử dụng cho tham chiếu dài hạn", mà không phải là ảnh tham chiếu liên lớp (ILRP);

dự báo một hoặc nhiều hình ảnh được lập mã trong CVS dựa trên giá trị của thành phần cú pháp.

Theo một phương án, sáng chế đề xuất phương pháp giải mã dòng bit video được lập mã, phương pháp này bao gồm:

xác định xem liệu điều kiện có được thỏa mãn hay không, trong đó điều kiện

này bao gồm chỉ số lớp của lớp hiện thời lớn hơn giá trị thiết đặt trước;

Khi điều kiện này được thỏa mãn, thì phân tích cú pháp thành phần cú pháp thứ nhất xác định liệu ít nhất một hình ảnh tham chiếu liên lớp (ILRP) có được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong chuỗi video được mã hóa (CVS) hay không;

dự báo một hoặc nhiều hình ảnh được lập mã trong CVS dựa trên giá trị của thành phần cú pháp thứ nhất.

Theo một phương án, trong đó giá trị thiết đặt trước bằng 0.

Theo một phương án, trong đó điều kiện còn bao gồm thành phần cú pháp thứ hai (ví dụ, `sps_video_parameter_set_id`) lớn hơn 0.

Theo một phương án, sáng chế đề xuất phương pháp giải mã dòng bit video được lập mã, phương pháp này bao gồm:

xác định xem liệu điều kiện có được thỏa mãn hay không, trong đó điều kiện này bao gồm chỉ số lớp của lớp hiện thời lớn hơn giá trị thiết đặt trước và mục nhập hiện thời trong cấu trúc danh sách hình ảnh tham chiếu là mục nhập ILRP;

khi điều kiện này được thỏa mãn, thì phân tích cú pháp thành phần cú pháp xác định chỉ số đối với danh sách các lớp phụ thuộc trực tiếp của lớp hiện thời;

dự báo một hoặc nhiều hình ảnh được lập mã trong CVS dựa trên cấu trúc danh sách hình ảnh tham chiếu mục nhập hiện thời mà ILRP thu được bằng cách sử dụng chỉ số đối với danh sách của lớp phụ thuộc trực tiếp.

Theo một phương án, trong đó giá trị thiết đặt trước bằng 1.

Theo một phương án, sáng chế đề xuất bộ mã hóa (20), bao gồm hệ mạch xử lý dùng để thực hiện phương pháp theo bất kỳ một trong số các phương án ở trên.

Theo một phương án, sáng chế đề xuất bộ giải mã (30), bao gồm hệ mạch xử lý dùng để thực hiện phương pháp theo bất kỳ một trong số các phương án ở trên.

Theo một phương án, sáng chế đề xuất sản phẩm chương trình máy tính, bao gồm mã chương trình dùng để thực hiện phương pháp theo bất kỳ một trong số các phương án ở trên khi được thực hiện trên máy tính hoặc bộ xử lý.

Theo một phương án, sáng chế đề xuất bộ giải mã, bao gồm:

một hoặc nhiều bộ xử lý; và

phương tiện lưu trữ bất biến có thể đọc được bằng máy tính được ghép nối với các bộ xử lý và lưu trữ chương trình dùng để được thực hiện bởi các bộ xử lý, trong đó chương trình này, khi được các bộ xử lý thực hiện, sẽ tạo cấu hình bộ

giải mã để thực hiện phương pháp theo bất kỳ một trong số các phương án ở trên.

Theo một phương án, sáng chế đề xuất bộ mã hóa, bao gồm:

một hoặc nhiều bộ xử lý; và

phương tiện lưu trữ bất biến có thể đọc được bằng máy tính được ghép nối với các bộ xử lý và lưu trữ chương trình dùng để được thực hiện bởi các bộ xử lý, trong đó chương trình này, khi được các bộ xử lý thực hiện, sẽ tạo cấu hình bộ mã hóa để thực hiện phương pháp theo bất kỳ một trong số các phương án ở trên.

Theo một phương án, sáng chế đề xuất phương tiện bất biến có thể đọc được bằng máy tính, mang mã chương trình which, khi được thiết bị máy tính thực hiện, sẽ khiến thiết bị máy tính này thực hiện phương pháp trong số bất kỳ một trong số các phương án ở trên.

Mô tả vắn tắt các hình vẽ

Các phương án sau đây của sáng chế sẽ được mô tả chi tiết hơn dựa vào các sơ đồ và hình vẽ đi kèm, trong đó:

FIG.1A là sơ đồ khối thể hiện ví dụ về hệ thống lập mã video được tạo cấu hình để thực hiện các phương án của sáng chế;

FIG.1B là sơ đồ khối thể hiện ví dụ khác về hệ thống lập mã video được tạo cấu hình để thực hiện các phương án của sáng chế;

FIG.2 là sơ đồ khối thể hiện ví dụ về bộ mã hóa video được tạo cấu hình để thực hiện các phương án của sáng chế;

FIG.3 là sơ đồ khối thể hiện cấu trúc được lấy làm ví dụ của bộ giải mã video được tạo cấu hình để thực hiện các phương án của sáng chế;

FIG.4 là sơ đồ khối minh họa ví dụ về thiết bị mã hóa hoặc thiết bị giải mã;

FIG.5 là sơ đồ khối minh họa ví dụ khác về thiết bị mã hóa hoặc thiết bị giải mã;

FIG.6 là sơ đồ khối thể hiện mã hóa mở rộng với 2 lớp;

FIG.7 là sơ đồ khối thể hiện cấu trúc được lấy làm ví dụ của hệ thống cung cấp nội dung 3100 mà nhận biết dịch vụ phân bổ nội dung.

FIG.8 là sơ đồ khối thể hiện cấu trúc được lấy làm ví dụ về thiết bị đầu cuối.

FIG.9 là lưu đồ của phương pháp giải mã theo một phương án.

FIG.10 là lưu đồ của phương pháp mã hóa theo một phương án.

FIG.11 là sơ đồ giản lược của bộ mã hóa theo một phương án.

FIG.12 là sơ đồ giản lược của bộ giải mã theo một phương án.

Các số chỉ dẫn giống nhau sau đây đề cập đến các đặc điểm giống hoặc ít nhất là tương đương về mặt chức năng nếu không được quy định khác.

Mô tả chi tiết sáng chế

Trong phần mô tả sau đây, thực hiện tham chiếu đến các hình vẽ đi kèm, mà tạo ra một phần của sáng chế, và thể hiện, bằng cách minh họa, các khía cạnh cụ thể của các phương án của sáng chế hoặc các khía cạnh cụ thể trong đó các phương án của sáng chế có thể được sử dụng. Cần hiểu rằng các phương án của sáng chế có thể được sử dụng trong khía cạnh khác và bao gồm các thay đổi về cấu trúc hoặc logic không được mô tả trong các hình vẽ. Do đó, phần mô tả chi tiết sau đây không được hiểu theo nghĩa giới hạn, và phạm vi của sáng chế được xác định bởi bộ yêu cầu bảo hộ đi kèm.

Ví dụ, cần hiểu rằng việc bộc lộ liên quan đến phương pháp được mô tả cũng có thể đúng với thiết bị hoặc hệ thống tương ứng được tạo cấu hình để thực hiện phương pháp này và ngược lại. Ví dụ, nếu một hoặc nhiều bước của phương pháp cụ thể được mô tả, thì thiết bị tương ứng có thể bao gồm một hoặc nhiều bộ phận, ví dụ bộ phận chức năng, để thực hiện một hoặc nhiều bước của phương pháp được mô tả này (ví dụ một bộ phận thực hiện một hoặc nhiều bước, hoặc nhiều bộ phận mỗi bộ phận thực hiện một hoặc nhiều bước), cho dù là một hoặc nhiều bộ phận như vậy không được mô tả hoặc được minh họa rõ ràng trong các hình vẽ. Mặt khác, ví dụ, nếu thiết bị cụ thể được mô tả dựa trên một hoặc nhiều bộ phận, ví dụ bộ phận chức năng, thì phương pháp tương ứng có thể bao gồm một bước để thực hiện chức năng của một hoặc nhiều bộ phận (ví dụ một bước thực hiện chức năng một hoặc nhiều bộ phận, hoặc nhiều bước mỗi bước thực hiện chức năng của một hoặc nhiều trong số các bộ phận), cho dù là một hoặc nhiều bước như vậy không được mô tả hoặc được minh họa rõ ràng trong các hình vẽ. Ngoài ra, cần hiểu rằng các đặc điểm của các phương án được lấy làm ví dụ khác và/hoặc các khía cạnh được mô tả ở đây có thể được kết hợp với nhau, trừ khi được xác định cụ thể khác.

Mã hóa video thông thường đề cập đến việc xử lý chuỗi các hình ảnh, mà tạo ra video hoặc chuỗi video. Thay vì thuật ngữ “hình ảnh” thuật ngữ “khung” hoặc “ảnh” có thể được sử dụng dưới dạng các từ đồng nghĩa trong lĩnh vực mã hóa video. Mã hóa video (hoặc mã hóa nói chung) bao gồm hai phần là mã hóa video và giải mã video. Mã hóa video được thực hiện ở phía nguồn, thông thường bao gồm việc xử lý (ví dụ bằng cách nén) các hình ảnh video gốc để giảm lượng dữ liệu cần có để thể hiện các hình ảnh video (dùng để lưu trữ và/hoặc truyền hiệu

quả hơn). Giải mã video được thực hiện ở phía đích và thông thường bao gồm xử lý ngược so với bộ mã hóa để tái cấu trúc các hình ảnh video. Các phương án đề cập đến “mã hóa” của các hình ảnh video (hoặc các hình ảnh nói chung) sẽ được hiểu là đề cập đến “mã hóa” hoặc “giải mã” các hình ảnh video hoặc các chuỗi video tương ứng. Sự kết hợp giữa bộ phận mã hóa và bộ phận giải mã còn được gọi là bộ mã hóa và giải mã (mã hóa và giải mã).

Trong trường hợp mã hóa video không suy hao, các hình ảnh video gốc có thể được tái cấu trúc, tức là các hình ảnh video được tái cấu trúc có cùng chất lượng với các hình ảnh video gốc (giả định không có suy hao truyền hoặc suy hao dữ liệu khác khi lưu trữ hoặc truyền). Trong trường hợp mã hóa video có suy hao, việc nén thêm, ví dụ bằng cách lượng tử hóa, được thực hiện, để giảm lượng dữ liệu thể hiện các hình ảnh video, điều này không thể được tái cấu trúc hoàn toàn ở bộ giải mã, tức là chất lượng của các hình ảnh video được tái cấu trúc thấp hơn hoặc kém hơn so với chất lượng của các hình ảnh video gốc.

Một vài tiêu chuẩn mã hóa video thuộc về nhóm “các bộ mã hóa và giải mã video lai suy hao” (tức là kết hợp dự báo không gian và thời gian trong cùng một miền và mã hóa biến đổi 2D (two dimension – hai chiều) để áp dụng lượng tử hóa trong miền biến đổi). Mỗi hình ảnh của chuỗi video thông thường được phân chia thành tập các khối không chồng lán và việc mã hóa thông thường được thực hiện trên mức khối. Nói cách khác, ở bộ mã hóa video thông thường được xử lý, tức là được mã hóa, trên mức khối (khối video), ví dụ bằng cách sử dụng dự báo không gian (nội ảnh) và/hoặc dự báo thời gian (liên ảnh) để tạo ra khối dự báo, trừ khối dự báo này khỏi khối hiện thời (khối hiện thời được xử lý/sẽ được xử lý) để thu khối dư, biến đổi khối dư và lượng tử hóa khối dư này trong miền biến đổi để giảm lượng dữ liệu sẽ được truyền (nén), trong khi ở bộ giải mã xử lý ngược so với bộ mã hóa được áp dụng cho khối được mã hóa hoặc được nén để tái cấu trúc khối hiện thời dùng cho việc thể hiện. Ngoài ra, bộ mã hóa sao lục vòng xử lý ở bộ giải mã sao cho cả hai sẽ tạo ra các dự báo (ví dụ các dự báo nội - liên ảnh) và/hoặc cấu trúc dùng giống hệt dùng cho việc xử lý, tức là mã hóa, các khối kế tiếp.

Các phương án sau đây về hệ thống lập mã video 10, bộ mã hóa video 20 và bộ giải mã video 30 được mô tả dựa vào Fig.1 đến Fig.3.

Fig.1A là sơ đồ khối giản lược minh họa hệ thống lập mã 10 được lấy làm ví dụ, ví dụ hệ thống lập mã video 10 (hoặc ngắn gọn là hệ thống lập mã 10) mà có thể sử dụng các kỹ thuật của sáng chế này. Bộ mã hóa video 20 (hoặc ngắn gọn là bộ mã hóa 20) và bộ giải mã video 30 (hoặc ngắn gọn là bộ giải mã 30) của hệ thống

lập mã video 10 thể hiện các ví dụ về các thiết bị mà có thể được tạo cấu hình để thực hiện các kỹ thuật theo các ví dụ khác nhau được mô tả trong sáng chế.

Như được thể hiện trên FIG.1A, hệ thống lập mã 10 bao gồm thiết bị nguồn 12 được tạo cấu hình để cung cấp dữ liệu hình ảnh được mã hóa 21 ví dụ đến thiết bị đích 14 để giải mã dữ liệu hình ảnh được mã hóa 13.

Thiết bị nguồn 12 bao gồm bộ mã hóa 20, và ngoài ra có thể, tức là một cách tùy chọn, bao gồm nguồn hình ảnh 16, bộ tiền xử lý (hoặc bộ phận tiền xử lý) 18, ví dụ bộ tiền xử lý hình ảnh 18, và giao diện truyền thông hoặc bộ truyền thông 22.

Nguồn hình ảnh 16 có thể bao gồm hoặc là loại bất kỳ trong số thiết bị chụp hình, ví dụ camera dùng để chụp hình ảnh thế giới thực và/hoặc loại bất kỳ trong số thiết bị tạo hình, ví dụ bộ xử lý đồ họa máy tính dùng để tạo ra hoạt ảnh máy tính, hoặc loại bất kỳ trong số thiết bị dùng để thu và/hoặc cung cấp hình ảnh thế giới thực hình ảnh được tạo ra bởi máy tính (ví dụ nội dung màn hình, hình ảnh thực tế ảo (virtual reality - VR)) và/hoặc bất kỳ kết hợp nào của chúng (ví dụ hình ảnh tương tác thực tế ảo (augmented reality - AR)). Nguồn hình ảnh có thể là loại bất kỳ trong số bộ nhớ hoặc bộ lưu trữ lưu trữ bất kỳ trong số các hình ảnh ở trên.

Để phân biệt với bộ tiền xử lý 18 và việc xử lý được thực hiện bởi bộ phận tiền xử lý 18, hình ảnh hoặc dữ liệu hình ảnh 17 có thể còn được gọi là hình ảnh thô hoặc dữ liệu hình ảnh thô 17.

Bộ tiền xử lý 18 được tạo cấu hình để thu nhận dữ liệu hình ảnh (thô) 17 và thực hiện việc tiền xử lý trên dữ liệu hình ảnh 17 để thu hình ảnh được tiền xử lý 19 hoặc dữ liệu hình ảnh được tiền xử lý 19. Việc tiền xử lý được thực hiện bởi bộ tiền xử lý 18 có thể, ví dụ, bao gồm cắt, chuyển đổi định dạng màu (ví dụ từ RGB thành YCbCr), hiệu chỉnh màu, hoặc giảm nhiễu. Cần hiểu rằng bộ phận tiền xử lý 18 có thể là bộ phận tùy chọn.

Bộ mã hóa video 20 được tạo cấu hình để thu nhận dữ liệu hình ảnh được tiền xử lý 19 và cung cấp dữ liệu hình ảnh được mã hóa 21 (chi tiết sẽ còn được mô tả dưới đây, ví dụ, dựa trên Fig.2).

Giao diện truyền thông 22 của thiết bị nguồn 12 có thể được tạo cấu hình để thu nhận dữ liệu hình ảnh được mã hóa 21 và truyền dữ liệu hình ảnh được mã hóa 21 (hoặc bất kỳ phiên bản được xử lý thêm nào của nó) qua kênh truyền thông 13 đến thiết bị khác, ví dụ thiết bị đích 14 hoặc bất kỳ thiết bị khác, để lưu trữ hoặc tái cấu trúc trực tiếp.

Thiết bị đích 14 bao gồm bộ giải mã 30 (ví dụ bộ giải mã video 30), và có thể còn, tức là một cách tùy chọn, bao gồm giao diện truyền thông hoặc bộ truyền thông 28, bộ xử lý sau 32 (hoặc bộ phận xử lý sau 32) và thiết bị hiển thị 34.

Giao diện truyền thông 28 của thiết bị đích 14 được tạo cấu hình để thu nhận dữ liệu hình ảnh được mã hóa 21 (hoặc bất kỳ phiên bản được xử lý thêm nào của nó), ví dụ một cách trực tiếp từ thiết bị nguồn 12 hoặc từ bất kỳ nguồn nào khác, ví dụ thiết bị lưu trữ, ví dụ thiết bị lưu trữ dữ liệu hình ảnh được mã hóa, và cung cấp dữ liệu hình ảnh được mã hóa 21 đến bộ giải mã 30.

Giao diện truyền thông 22 và giao diện truyền thông 28 có thể được tạo cấu hình để truyền hoặc thu nhận dữ liệu hình ảnh được mã hóa 21 hoặc dữ liệu được mã hóa 13 qua liên kết truyền thông trực tiếp giữa thiết bị nguồn 12 và thiết bị đích 14, ví dụ qua kết nối không dây hoặc có dây trực tiếp, hoặc qua bất kỳ loại mạng nào, ví dụ mạng không dây hoặc có dây hoặc bất kỳ kết hợp nào của chúng, hoặc bất kỳ loại mạng riêng hoặc chung nào, hoặc bất kỳ loại kết hợp nào của chúng.

Giao diện truyền thông 22 có thể, ví dụ, được tạo cấu hình để gói dữ liệu hình ảnh được mã hóa 21 trong định dạng thích hợp, ví dụ gói, và/hoặc xử lý dữ liệu hình ảnh được mã hóa bằng cách sử dụng loại bất kỳ trong số mã hóa hoặc xử lý truyền để truyền qua liên kết truyền thông hoặc mạng truyền thông.

Giao diện truyền thông 28, tạo ra bản sao của giao diện truyền thông 22, có thể, ví dụ, được tạo cấu hình để thu nhận dữ liệu được truyền và xử lý dữ liệu truyền này bằng cách sử dụng loại bất kỳ trong số giải mã truyền tương ứng hoặc xử lý và/hoặc giải gói để thu dữ liệu hình ảnh được mã hóa 21.

Cả hai, giao diện truyền thông 22 và giao diện truyền thông 28 có thể được tạo cấu hình dưới dạng các giao diện truyền thông đơn hướng như được thể hiện bằng mũi tên dùng cho kênh truyền thông 13 trên Fig.1A chỉ hướng từ thiết bị nguồn 12 đến thiết bị đích 14, hoặc các giao diện truyền thông lưỡng hướng, và có thể được tạo cấu hình, ví dụ để gửi và thu nhận các bản tin, ví dụ để thiết lập kết nối, để xác nhận và trao đổi bất kỳ thông tin khác liên quan đến liên kết truyền thông và/hoặc dữ liệu truyền, ví dụ truyền dữ liệu hình ảnh được lập mã.

Bộ giải mã 30 được tạo cấu hình để thu nhận dữ liệu hình ảnh được mã hóa 21 và cung cấp dữ liệu hình ảnh được giải mã 31 hoặc hình ảnh được giải mã 31 (chi tiết sẽ còn được mô tả dưới đây, ví dụ, dựa vào Fig.3 hoặc Fig.5).

Bộ xử lý sau 32 của thiết bị đích 14 được tạo cấu hình để xử lý sau dữ liệu hình ảnh được giải mã 31 (còn được gọi là dữ liệu hình ảnh được tái cấu trúc), ví dụ hình ảnh được giải mã 31, để thu dữ liệu hình ảnh được xử lý sau 33, ví dụ hình

ảnh được xử lý sau 33. Việc xử lý sau được thực hiện bởi bộ phận xử lý sau 32 có thể bao gồm, ví dụ chuyển đổi định dạng màu (ví dụ từ YCbCr thành RGB), hiệu chỉnh màu, cắt, hoặc lấy mẫu lại, hoặc bất kỳ xử lý nào khác, ví dụ để chuẩn bị dữ liệu hình ảnh được giải mã 31 cho việc hiển thị, ví dụ bởi thiết bị hiển thị 34.

Thiết bị hiển thị 34 của thiết bị đích 14 được tạo cấu hình để thu nhận dữ liệu hình ảnh được xử lý sau 33 để hiển thị hình ảnh này, ví dụ đến người dùng hoặc người xem. Thiết bị hiển thị 34 có thể hoặc bao gồm loại màn hình bất kỳ để thể hiện hình ảnh được tái cấu trúc, ví dụ màn hình tích hợp hoặc màn hình bên ngoài hoặc màn hình. Các màn hình này có thể, ví dụ bao gồm các màn hình tinh thể lỏng (liquid crystal display - LCD), màn hình điôt phát quang hữu cơ (organic light emitting diode - OLED), các màn hình plasma, các máy chiếu, các màn hình micro LED (light emitting diode - điôt phát quang), tinh thể lỏng trên silic (liquid crystal on silicon - LCoS), bộ xử lý ánh sáng kỹ thuật số (DLP - digital light processor) hoặc bất kỳ loại màn hình khác nào.

Mặc dù Fig.1A mô tả thiết bị nguồn 12 và thiết bị đích 14 dưới dạng các thiết bị riêng rẽ, nhưng các phương án về các thiết bị có thể còn bao gồm cả hai hoặc hai chức năng, thiết bị nguồn 12 hoặc chức năng tương ứng và thiết bị đích 14 hoặc chức năng tương ứng. Trong các phương án như vậy thiết bị nguồn 12 hoặc chức năng tương ứng và thiết bị đích 14 hoặc chức năng tương ứng có thể được thực hiện bằng cách sử dụng cùng một phần cứng và/hoặc phần mềm hoặc bằng phần cứng và/hoặc phần mềm riêng rẽ hoặc bất kỳ kết hợp nào của chúng.

Như sẽ được người có trình độ trung bình trong lĩnh vực hiểu rõ dựa trên phần mô tả, sự tồn tại và phân chia (chính xác) các chức năng của các bộ phận khác nhau hoặc các chức năng có trong thiết bị nguồn 12 và/hoặc thiết bị đích 14 như được thể hiện trên FIG.1A có thể thay đổi tùy thuộc vào thiết bị và ứng dụng thực tế.

Bộ mã hóa **20** (ví dụ bộ mã hóa video 20) hoặc bộ giải mã **30** (ví dụ bộ giải mã video 30) hoặc cả bộ mã hóa 20 và bộ giải mã 30 có thể được thực hiện qua hệ mạch xử lý như được thể hiện trên FIG.1B, chẳng hạn như một hoặc nhiều bộ vi xử lý, các bộ xử lý tín hiệu kỹ thuật số (DSPs - digital signal processors), các mạch tích hợp chuyên dụng (application-specific integrated circuits - (ASICs), các mảng cổng lập trình được dạng trường (field-programmable gate arrays - FPGAs), logic rời rạc, phần cứng, mã hóa video dành riêng hoặc bất kỳ các kết hợp nào của chúng. Bộ mã hóa 20 có thể được thực hiện qua hệ mạch xử lý 46 để bao gồm các môđun khác nhau như được thảo luận đối với bộ mã hóa 20 của FIG.2 và/hoặc bất kỳ hệ thống lập mã hoặc hệ thống con nào khác được mô tả ở

đây. Bộ giải mã 30 có thể được thực hiện qua hệ mạch xử lý 46 đề bao gồm các môđun khác nhau như được thảo luận đối với bộ giải mã 30 của FIG.3 và/hoặc bất kỳ hệ thống giải mã hoặc hệ thống con nào khác được mô tả ở đây. Hệ mạch xử lý có thể được tạo cấu hình để thực hiện các hoạt động khác nhau như được thảo luận sau. Như được thể hiện trên FIG.5, nếu các kỹ thuật được thực hiện một phần trong phần mềm, thì thiết bị có thể lưu trữ các lệnh dùng cho phần mềm này trong phương tiện lưu trữ có thể đọc được bằng máy tính thích hợp, bất biến và có thể thực hiện các lệnh trong phần cứng bằng cách sử dụng một hoặc nhiều bộ xử lý để thực hiện các kỹ thuật của sáng chế này. Một trong hai bộ mã hóa video **20** và bộ giải mã video **30** có thể được tích hợp dưới dạng một phần của bộ mã hóa/bộ giải mã kết hợp (bộ mã hóa và giải mã) trong một thiết bị, ví dụ, như được thể hiện trên FIG.1B.

Thiết bị nguồn 12 và thiết bị đích 14 có thể bao gồm bất kỳ trong số một loạt các thiết bị, bao gồm loại bất kỳ trong số các thiết bị cầm tay hoặc tĩnh, ví dụ notebook hoặc các máy tính xách tay, các điện thoại di động, các điện thoại thông minh, các thiết bị bảng hoặc các máy tính bảng, các camera, các máy tính để bàn, các đầu thu giải mã tín hiệu, tivi, các thiết bị hiển thị, các máy nghe nhạc kỹ thuật số, máy chơi trò chơi video, các thiết bị truyền trực tuyến video (chẳng hạn như các máy chủ dịch vụ nội dung hoặc các máy chủ phân bổ nội dung), thiết bị thu nhận phát quảng bá, thiết bị truyền phát quảng bá, hoặc tương tự và có thể không sử dụng hoặc loại hệ điều hành bất kỳ. Trong một vài trường hợp, thiết bị nguồn 12 và thiết bị đích 14 có thể được trang bị cho truyền thông không dây. Do đó, thiết bị nguồn 12 và thiết bị đích 14 có thể là các thiết bị truyền thông không dây.

Trong một vài trường hợp, hệ thống lập mã video **10** được minh họa trên Fig.1A chỉ là ví dụ và các kỹ thuật theo sáng chế có thể áp dụng cho các thiết bị mã hóa video (ví dụ, mã hóa video hoặc giải mã video) mà không nhất thiết phải bao gồm bất kỳ truyền thông dữ liệu nào giữa các thiết bị mã hóa và giải mã. Trong các ví dụ khác, dữ liệu được truy lục từ bộ nhớ cục bộ, được truyền trực tuyến qua mạng, hoặc tương tự. Thiết bị mã hóa video có thể mã hóa và lưu trữ dữ liệu vào bộ nhớ, và/hoặc thiết bị giải mã video có thể truy lục và giải mã dữ liệu này từ bộ nhớ. Trong một vài ví dụ, việc mã hóa và giải mã được thực hiện bởi các thiết bị mà không truyền thông với nhau, nhưng mã hóa đơn giản dữ liệu vào bộ nhớ và/hoặc truy lục và giải mã dữ liệu từ bộ nhớ này.

Để thuận tiện cho việc mô tả, các phương án của sáng chế được mô tả ở đây, ví dụ, bằng cách tham chiếu đến mã hóa video hiệu quả cao (HEVC, High-Efficiency

Video Coding) hoặc đến phần mềm tham chiếu của mã hóa video đa năng (VVC - Versatile Video coding), tiêu chuẩn mã hóa video thế hệ kế tiếp được phát triển bởi nhóm hợp tác liên kết mã hóa video (JCT-VC, Joint Collaboration Team on Video Coding) của nhóm chuyên gia mã hóa video (VCEG - Video Coding Experts Group) ITU-T và nhóm chuyên gia hình ảnh động (MPEG - Motion Picture Experts Group) ISO/IEC. Một trong số người có trình độ trung trình trong lĩnh vực kỹ thuật sẽ hiểu rằng các phương án của sáng chế sẽ không bị giới hạn ở HEVC hoặc VVC.

Bộ mã hóa và phương pháp mã hóa

Fig.2 thể hiện sơ đồ khối giản lược về bộ mã hóa video 20 được lấy làm ví dụ mà được tạo cấu hình để thực hiện các kỹ thuật theo sáng chế. Trong ví dụ của FIG.2, bộ mã hóa video 20 bao gồm đầu ra 201 (hoặc giao diện đầu ra 201), bộ tính toán phần dư 204, bộ xử lý biến đổi 206, bộ lượng tử hóa 208, bộ lượng tử hóa ngược 210, và bộ xử lý biến đổi ngược 212, bộ tái cấu trúc 214, bộ lọc vòng 220, bộ đệm hình ảnh được giải mã (DPB) 230, bộ lựa chọn chế độ 260, bộ mã hóa entropi 270 và đầu ra 272 (hoặc giao diện đầu ra 272). Bộ lựa chọn chế độ 260 có thể bao gồm bộ dự báo liên ảnh 244, bộ dự báo nội ảnh 254 và bộ phân chia 262. Bộ dự báo liên ảnh 244 có thể bao gồm bộ ước lượng chuyển động và bộ bù chuyển động (không được thể hiện). Bộ mã hóa video 20 như được thể hiện trên FIG.2 có thể còn được gọi là bộ mã hóa video lai hoặc bộ mã hóa video theo bộ mã hóa và giải mã video lai.

Bộ tính toán phần dư 204, bộ xử lý biến đổi 206, bộ lượng tử hóa 208, bộ lựa chọn chế độ 260 có thể được gọi là tạo ra đường dẫn tín hiệu thuận của bộ mã hóa 20, trong khi bộ lượng tử hóa ngược 210, bộ xử lý biến đổi ngược 212, bộ tái cấu trúc 214, bộ đệm 216, bộ lọc vòng 220, bộ đệm hình ảnh được giải mã (DPB) 230, bộ dự báo liên ảnh 244 và bộ dự báo nội ảnh 254 có thể được gọi là tạo ra đường dẫn tín hiệu ngược của bộ mã hóa video 20, trong đó đường dẫn tín hiệu ngược của bộ mã hóa video 20 tương ứng với đường dẫn tín hiệu của bộ giải mã (xem bộ giải mã video 30 trên Fig.3). Bộ lượng tử hóa ngược 210, bộ xử lý biến đổi ngược 212, bộ tái cấu trúc 214, bộ lọc vòng 220, bộ đệm hình ảnh được giải mã (DPB) 230, bộ dự báo liên ảnh 244 và bộ dự báo nội ảnh 254 còn được gọi là tạo ra “bộ giải mã tích hợp” của bộ mã hóa video 20.

Các hình ảnh & phân chia hình ảnh (các hình ảnh & các khối)

Bộ mã hóa 20 có thể được tạo cấu hình để thu nhận, ví dụ qua đầu ra 201, hình ảnh 17 (hoặc dữ liệu hình ảnh 17), ví dụ hình ảnh gồm chuỗi các hình ảnh tạo ra

video hoặc chuỗi video. Hình ảnh hoặc dữ liệu hình ảnh được thu nhận có thể còn là hình ảnh được tiền xử lý 19 (hoặc dữ liệu hình ảnh được tiền xử lý 19). Để cho đơn giản phần mô tả sau đây đề cập đến hình ảnh 17. Hình ảnh 17 có thể còn được gọi là hình ảnh hiện thời hoặc hình ảnh sẽ được lập mã (cụ thể là khi mã hóa video để phân biệt hình ảnh hiện thời với các hình ảnh khác, ví dụ các hình ảnh được mã hóa và/hoặc được giải mã trước đó của cùng một chuỗi video, tức là chuỗi video mà cũng bao gồm hình ảnh hiện thời).

Hình ảnh (kỹ thuật số) là hoặc có thể được coi là mảng hai chiều hoặc ma trận của các mẫu có các giá trị cường độ. Mẫu trong mảng này có thể còn được gọi là điểm ảnh (dạng ngắn gọn của thành phần điểm ảnh) hoặc pel. Số lượng các mẫu theo hướng ngang hoặc dọc (hoặc trục) của mảng hoặc hình ảnh xác định kích cỡ và/hoặc độ phân giải của hình ảnh này. Để thể hiện màu sắc, thông thường ba thành phần màu được sử dụng, tức là hình ảnh có thể được thể hiện hoặc bao gồm ba mảng mẫu. Trong định dạng RGB hoặc không gian màu hình ảnh bao gồm mảng mẫu màu đỏ, màu xanh lá, màu xanh da trời tương ứng. Tuy nhiên, khi mã hóa video mỗi điểm ảnh thông thường được thể hiện ở định dạng độ chói và sắc độ hoặc không gian màu, ví dụ YCbCr, mà bao gồm thành phần độ chói được thể hiện bằng Y (đôi khi L được cũng sử dụng thay thế) và hai thành phần sắc độ được thể hiện bằng Cb và Cr. Thành phần độ chói (hoặc ngắn gọn là luma) Y thể hiện độ sáng hoặc cường độ mức màu xám (ví dụ giống như hình ảnh trên thang màu xám), trong khi hai thành phần sắc độ (hoặc ngắn gọn là chroma) Cb và Cr thể hiện độ màu hoặc các thành phần thông tin màu sắc. Do đó, hình ảnh ở định dạng YCbCr bao gồm mảng mẫu độ chói có các giá trị mẫu độ chói (Y), và hai mảng mẫu sắc độ có các giá trị sắc độ (Cb và Cr). Các hình ảnh ở định dạng RGB có thể được chuyển đổi hoặc được biến đổi thành định dạng YCbCr và ngược lại, việc xử lý này còn được biết đến là biến đổi hoặc chuyển đổi màu sắc. Nếu hình ảnh là đơn sắc, thì hình ảnh này có thể chỉ bao gồm mảng mẫu độ chói. Do đó, hình ảnh này có thể là, ví dụ, mảng có các mẫu độ chói ở định dạng đơn sắc hoặc mảng có các mẫu độ chói và hai mảng có các mẫu sắc độ tương ứng ở định dạng màu 4:2:0, 4:2:2, và 4:4:4.

Các phương án về bộ mã hóa video 20 có thể bao gồm bộ phân chia hình ảnh (không được mô tả trên Fig.2) được tạo cấu hình để phân chia hình ảnh 17 thành nhiều (thông thường không chồng lên nhau) khối hình ảnh 203. Các khối này có thể còn được gọi là các khối gốc, các khối macro (H.264/AVC) hoặc khối cây mã hóa (CTB - coding tree block) hoặc các đơn vị cây mã hóa (CTU - coding tree unit) (H.265/HEVC và VVC). Bộ phân chia hình ảnh có thể được tạo cấu hình để

sử dụng cùng một kích cỡ khối cho tất cả các hình ảnh của chuỗi video và lưới tương ứng xác định kích cỡ khối, hoặc thay đổi kích cỡ khối giữa các hình ảnh hoặc các tập con hoặc các nhóm của các hình ảnh, và phân chia mỗi hình ảnh thành các khối tương ứng.

Trong các phương án khác, bộ mã hóa video có thể được tạo cấu hình để thu nhận một cách trực tiếp khối 203 của hình ảnh 17, ví dụ một, một vài hoặc tất cả các khối tạo ra hình ảnh 17. Khối hình ảnh 203 có thể còn được gọi là khối hình ảnh hiện thời hoặc khối hình ảnh sẽ được lập mã.

Giống như hình ảnh 17, khối hình ảnh 203 lại là hoặc có thể được coi là mảng hai chiều hoặc ma trận của các mẫu có các giá trị cường độ (các giá trị mẫu), mặc dù có kích thước nhỏ hơn hình ảnh 17. Nói cách khác, khối 203 có thể bao gồm, ví dụ, một mảng mẫu (ví dụ mảng độ chói trong trường hợp của hình ảnh đơn sắc 17, hoặc mảng độ chói hoặc sắc độ trong trường hợp của hình ảnh màu) hoặc ba mảng mẫu (ví dụ mảng độ chói và hai mảng sắc độ trong trường hợp của hình ảnh màu 17) hoặc bất kỳ số lượng và/hoặc loại mảng nào khác tùy thuộc vào định dạng màu được áp dụng. Số lượng các mẫu theo hướng ngang hoặc dọc (hoặc trục) của khối 203 xác định kích cỡ của khối 203. Do đó, khối có thể, ví dụ, mảng $M \times N$ (M-cột nhân N-dòng) có các mẫu, hoặc mảng $M \times N$ có các hệ số biến đổi.

Các phương án về bộ mã hóa video 20 như được thể hiện trên FIG.2 có thể được tạo cấu hình để mã hóa hình ảnh 17 thành khối nhân khối, ví dụ việc mã hóa và dự báo được thực hiện trên mỗi khối 203.

Các phương án về bộ mã hóa video 20 như được thể hiện trên FIG.2 có thể còn được tạo cấu hình để phân chia và/hoặc mã hóa hình ảnh bằng cách sử dụng các lát (còn được gọi là các lát video), trong đó hình ảnh có thể được phân chia thành hoặc được mã hóa bằng cách sử dụng một hoặc nhiều lát (thông thường không chồng lên nhau), và mỗi lát có thể bao gồm một hoặc nhiều khối (ví dụ các CTU) hoặc một hoặc nhiều nhóm gồm các khối (ví dụ các ô (H.265/HEVC và VVC) hoặc các viên (VVC)).

Các phương án về bộ mã hóa video 20 như được thể hiện trên FIG.2 có thể còn được tạo cấu hình để phân chia thành và/hoặc mã hóa hình ảnh bằng cách sử dụng các lát/các nhóm ô (còn được gọi là các nhóm ô video) và/hoặc các ô (còn được gọi là các ô video), trong đó hình ảnh có thể được phân chia thành hoặc được mã hóa bằng cách sử dụng một hoặc nhiều lát/các nhóm ô (thông thường không chồng lên nhau), và mỗi lát/nhóm ô có thể bao gồm, ví dụ một hoặc nhiều khối (ví dụ các CTU) hoặc một hoặc nhiều ô, trong đó mỗi ô, ví dụ có thể có dạng hình chữ

nhật và có thể bao gồm một hoặc nhiều khối (ví dụ các CTU), ví dụ các khối hoàn chỉnh hoặc phân đoạn.

Tính toán phần dư

Bộ tính toán phần dư 204 có thể được tạo cấu hình để tính toán khối dư 205 (còn được gọi là phần dư 205) dựa trên khối hình ảnh 203 và khối dự báo 265 (chi tiết thêm về khối dự báo 265 sẽ được cung cấp sau), ví dụ bằng cách trừ các giá trị mẫu của khối dự báo 265 khỏi các giá trị mẫu của khối hình ảnh 203, mẫu nhân mẫu (điểm ảnh nhân điểm ảnh) để thu khối dư 205 trong cùng một miền.

Biến đổi

Bộ xử lý biến đổi 206 có thể được tạo cấu hình để áp dụng biến đổi, ví dụ biến đổi cosin rời rạc (DCT - discrete cosine transform) hoặc biến đổi sin rời rạc (DST - discrete sine transform), trên các giá trị mẫu của khối dư 205 để thu các hệ số biến đổi 207 trong miền biến đổi. Các hệ số biến đổi 207 có thể còn được gọi là các hệ số biến đổi dư và thể hiện khối dư 205 trong miền biến đổi.

Bộ xử lý biến đổi 206 có thể được tạo cấu hình để áp dụng các phép lấy xấp xỉ số nguyên của DCT/DST, chẳng hạn như các biến đổi được xác định cho H.265/HEVC. So với biến đổi DCT trực giao, các phép lấy xấp xỉ số nguyên như vậy thông thường được chia tỷ lệ bởi hệ số nhất định. Để duy trì định chuẩn của khối dư mà được xử lý bằng các biến đổi thuận và ngược, các hệ số tỷ lệ bổ sung được áp dụng như một phần của xử lý biến đổi. Các hệ số tỷ lệ thông thường được chọn dựa trên các giới hạn nhất định như các hệ số tỷ lệ là lũy thừa của hai đối với các phép toán độ dịch, độ sâu bit của các hệ số biến đổi, sự đánh đổi giữa độ chính xác và các chi phí thực hiện, vân vân. Các hệ số tỷ lệ cụ thể, ví dụ, được xác định cho biến đổi ngược, ví dụ bởi bộ xử lý biến đổi ngược 212 (và biến đổi ngược tương ứng, ví dụ bởi bộ xử lý biến đổi ngược 312 ở bộ giải mã video 30) và các hệ số tỷ lệ tương ứng cho biến đổi thuận, ví dụ bởi bộ xử lý biến đổi 206, ở bộ mã hóa 20 do đó có thể được xác định.

Các phương án về bộ mã hóa video 20 (bộ xử lý biến đổi 206 tương ứng) có thể được tạo cấu hình để xuất ra các tham số biến đổi, ví dụ loại biến đổi hoặc các biến đổi, ví dụ một cách trực tiếp hoặc được mã hóa hoặc được nén qua bộ mã hóa entropi 270, sao cho, ví dụ, bộ giải mã video 30 có thể thu nhận và sử dụng các tham số biến đổi này cho việc giải mã.

Lượng tử hóa

Bộ lượng tử hóa 208 có thể được tạo cấu hình để lượng tử hóa các hệ số biến đổi 207 để thu các hệ số lượng tử hóa 209, ví dụ bằng cách áp dụng lượng tử hóa vô hướng hoặc lượng tử hóa vector. Các hệ số lượng tử hóa 209 có thể còn được gọi là các hệ số biến đổi được lượng tử hóa 209 hoặc các hệ số dư được lượng tử hóa 209.

Xử lý lượng tử hóa có thể giảm độ sâu bit liên kết với một vài hoặc tất cả các hệ số biến đổi 207. Ví dụ, hệ số biến đổi bit n có thể được làm tròn xuống thành biến đổi bit m khi lượng tử hóa, trong đó n lớn hơn m . Mức lượng tử hóa có thể được cải biến bằng cách điều chỉnh tham số lượng tử hóa (QP - quantization parameter). Ví dụ đối với lượng tử hóa vô hướng, việc chia tỷ lệ khác nhau có thể được áp dụng để thu được lượng tử hóa mịn hơn hoặc thô hơn. Các kích cỡ bước lượng tử hóa nhỏ hơn tương ứng với lượng tử hóa mịn hơn, trong khi các kích cỡ bước lượng tử hóa lớn hơn tương ứng với lượng tử hóa thô hơn. Kích cỡ bước lượng tử hóa có thể áp dụng có thể được thể hiện bằng tham số lượng tử hóa (QP). Tham số lượng tử hóa này có thể ví dụ là chỉ số đối với tập xác định trước có các kích cỡ bước lượng tử hóa có thể áp dụng. Ví dụ, các tham số lượng tử hóa nhỏ có thể tương ứng với lượng tử hóa mịn (các kích cỡ bước lượng tử hóa nhỏ) và các tham số lượng tử hóa lớn có thể tương ứng với lượng tử hóa thô (các kích cỡ bước lượng tử hóa lớn) hoặc ngược lại. Lượng tử hóa có thể bao gồm phân chia theo kích cỡ bước lượng tử hóa và giải lượng tử hóa tương ứng và/hoặc ngược, ví dụ bởi bộ lượng tử hóa ngược 210, có thể bao gồm phép nhân với kích cỡ bước lượng tử hóa. Các phương án theo một vài tiêu chuẩn, ví dụ HEVC, có thể được tạo cấu hình để sử dụng tham số lượng tử hóa để xác định kích cỡ bước lượng tử hóa. Thông thường, kích cỡ bước lượng tử hóa có thể được tính toán dựa trên tham số lượng tử hóa bằng cách sử dụng phép lấy xấp xỉ điểm cố định của phương trình có phép chia. Các hệ số tỷ lệ bổ sung có thể được đưa vào lượng tử hóa và giải lượng tử hóa để khôi phục định chuẩn của khối dư, mà có thể được cải biến do việc chia tỷ lệ được sử dụng trong phép lấy xấp xỉ điểm cố định của phương trình đối với kích cỡ bước lượng tử hóa và tham số lượng tử hóa. Trong một cách thức thực hiện ví dụ, việc chia tỷ lệ của biến đổi ngược và giải lượng tử hóa có thể được kết hợp. Ngoài ra, các bảng lượng tử hóa được tùy biến có thể được sử dụng và được báo hiệu từ bộ mã hóa đến bộ giải mã, ví dụ trong dòng bit. Lượng tử hóa là hoạt động làm suy hao, trong đó suy hao tăng cùng với sự gia tăng các kích cỡ bước lượng tử hóa.

Các phương án về bộ mã hóa video 20 (bộ lượng tử hóa 208 tương ứng) có thể được tạo cấu hình để xuất ra các tham số lượng tử hóa (QP), ví dụ một cách trực

tiếp hoặc được mã hóa qua bộ mã hóa entropi 270, sao cho, ví dụ, bộ giải mã video 30 có thể thu nhận và áp dụng các tham số lượng tử hóa cho việc giải mã.

Lượng tử hóa ngược

Bộ lượng tử hóa ngược 210 được tạo cấu hình để áp dụng lượng tử hóa ngược của bộ lượng tử hóa 208 trên các hệ số lượng tử hóa để thu các hệ số giải lượng tử hóa 211, ví dụ bằng cách áp dụng ngược sơ đồ lượng tử hóa được áp dụng bởi bộ lượng tử hóa 208 dựa trên hoặc bằng cách sử dụng kích cỡ bước lượng tử hóa giống như bộ lượng tử hóa 208. Các hệ số giải lượng tử hóa 211 có thể còn được gọi là các hệ số dư được giải lượng tử hóa 211 và tương ứng - mặc dù thông thường không giống với các hệ số biến đổi do suy hao bằng cách lượng tử hóa – với các hệ số biến đổi 207.

Biến đổi ngược

Bộ xử lý biến đổi ngược 212 được tạo cấu hình để áp dụng biến đổi ngược của biến đổi được áp dụng bởi bộ xử lý biến đổi 206, ví dụ biến đổi ngược cosin rời rạc (DCT) hoặc biến đổi ngược sin rời rạc (DST) hoặc các biến đổi khác, để thu khối dư được tái cấu trúc 213 (hoặc các hệ số giải lượng tử hóa 213 tương ứng) trong cùng một miền. Khối dư được tái cấu trúc 213 có thể còn được gọi là khối biến đổi 213.

Tái cấu trúc

Bộ tái cấu trúc 214 (ví dụ bộ cộng hoặc máy cộng 214) được tạo cấu hình để bổ sung khối biến đổi 213 (tức là khối dư được tái cấu trúc 213) vào khối dự báo 265 để thu khối được tái cấu trúc 215 trong cùng một miền, ví dụ bằng cách bổ sung – mẫu nhân mẫu - các giá trị mẫu của khối dư được tái cấu trúc 213 và các giá trị mẫu của khối dự báo 265.

Lọc

Bộ phận lọc vòng 220 (hoặc ngắn gọn là “bộ lọc vòng” 220), được tạo cấu hình để lọc khối được tái cấu trúc 215 để thu khối được lọc 221, hoặc nói chung, để lọc các mẫu được tái cấu trúc để thu các giá trị mẫu được lọc. Bộ lọc vòng, ví dụ, được tạo cấu hình để làm nhẵn các chuyển tiếp điểm ảnh, hoặc nếu không thì cải thiện chất lượng video. Bộ lọc vòng 220 có thể bao gồm một hoặc nhiều bộ lọc vòng chẳng hạn như bộ lọc giải khối, bộ lọc độ dịch tương thích mẫu (SAO, sample-adaptive offset) hoặc một hoặc nhiều bộ lọc khác, ví dụ bộ lọc vòng tương thích (ALF - adaptive loop filter), bộ lọc triệt tiêu nhiễu (NSF - noise suppression filter), hoặc bất kỳ kết hợp nào của chúng. Trong một ví dụ, bộ lọc vòng 220 có

thể bao gồm bộ lọc giải khối, bộ lọc SAO và bộ lọc ALF. Thứ tự xử lý lọc có thể là lọc giải khối, SAO và ALF. Trong một ví dụ khác, xử lý được gọi là ánh xạ độ chói có chia tỷ lệ sắc độ (luma mapping with chroma scaling (LMCS)) (tức là, bộ tái tạo hình dạng trong vòng tương thích) được bổ sung. Xử lý này được thực hiện trước khi giải khối. Trong một ví dụ khác, xử lý lọc giải khối có thể còn được áp dụng cho các mép của khối con bên trong, ví dụ các mép của các khối con afin, mép của các khối con ATMVP, các mép biến đổi khối con (SBT, sub-block transform) và các mép nội phân chia con (ISP, intra sub-partition). Mặc dù bộ lọc vòng 220 được thể hiện trên Fig.2 dưới dạng là bộ lọc trong vòng, nhưng trong các cấu hình khác, bộ lọc vòng 220 có thể được thực hiện dưới dạng bộ lọc vòng sau. Khối được lọc 221 có thể còn được gọi là khối được lọc được tái cấu trúc 221.

Các phương án về bộ mã hóa video 20 (bộ lọc vòng 220 tương ứng) có thể được tạo cấu hình để xuất ra các tham số bộ lọc vòng (chẳng hạn như các tham số bộ lọc SAO hoặc các tham số bộ lọc ALF hoặc các tham số LMCS), ví dụ một cách trực tiếp hoặc được mã hóa qua bộ mã hóa entropi 270, sao cho, ví dụ, bộ giải mã 30 có thể thu nhận và áp dụng các tham số bộ lọc vòng hoặc các bộ lọc vòng tương ứng giống nhau cho việc giải mã.

Bộ đệm hình ảnh được giải mã

Bộ đệm hình ảnh được giải mã (DPB) 230 có thể là bộ nhớ mà lưu trữ các hình ảnh tham chiếu, hoặc nói chung là dữ liệu hình ảnh tham chiếu, dùng để mã hóa dữ liệu video bởi bộ mã hóa video 20. DPB 230 có thể được tạo ra bởi bất kỳ trong số nhiều thiết bị bộ nhớ, chẳng hạn như bộ nhớ truy nhập ngẫu nhiên động (DRAM - dynamic random access), bao gồm DRAM đồng bộ (SDRAM, synchronous DRAM), RAM từ điện trở (MRAM, magnetoresistive RAM), RAM điện trở (RRAM - resistive RAM), hoặc các loại thiết bị bộ nhớ khác. Bộ đệm hình ảnh được giải mã (DPB) 230 có thể được tạo cấu hình để lưu trữ một hoặc nhiều khối được lọc 221. Bộ đệm hình ảnh được giải mã 230 có thể còn được tạo cấu hình để lưu trữ các khối được lọc trước đó khác, ví dụ các khối được tái cấu trúc và được lọc trước đó 221, của cùng một hình ảnh hiện thời hoặc của các hình ảnh khác nhau, ví dụ các hình ảnh được tái cấu trúc trước đó, và có thể cung cấp các hình ảnh được tái cấu trúc hoàn toàn trước đó, tức là được giải mã, (và các khối và các mẫu tham chiếu tương ứng) và/hoặc hình ảnh hiện thời được tái cấu trúc một phần (và các khối và các mẫu tham chiếu tương ứng), ví dụ dùng để dự báo liên ảnh. Bộ đệm hình ảnh được giải mã (DPB) 230 có thể còn được tạo cấu

hình để lưu trữ một hoặc nhiều khối được tái cấu trúc chưa được lọc 215, hoặc nói chung là các khối được tái cấu trúc chưa được lọc, ví dụ nếu khối được tái cấu trúc 215 không được lọc bởi bộ lọc vòng 220, hoặc bất kỳ phiên bản được xử lý thêm nào khác của các khối hoặc mẫu được tái cấu trúc.

Lựa chọn chế độ (phân chia và dự báo)

Bộ lựa chọn chế độ 260 bao gồm bộ phân chia 262, bộ dự báo nội ảnh 244 và bộ dự báo nội ảnh 254, và được tạo cấu hình để thu nhận hoặc thu dữ liệu hình ảnh gốc, ví dụ khối gốc 203 (khối hiện thời 203 của hình ảnh hiện thời 17), và dữ liệu hình ảnh được tái cấu trúc, ví dụ các mẫu hoặc khối được tái cấu trúc được lọc và/hoặc chưa được lọc của cùng một hình ảnh (hiện thời) và/hoặc từ một hoặc nhiều trong số các hình ảnh được giải mã trước đó, ví dụ từ bộ đệm hình ảnh được giải mã 230 hoặc các bộ đệm khác (ví dụ bộ đệm tuyến, không được thể hiện). Dữ liệu hình ảnh được tái cấu trúc được sử dụng làm dữ liệu hình ảnh tham chiếu dùng để dự báo, ví dụ dự báo liên ảnh hoặc dự báo nội ảnh, để thu khối dự báo 265 hoặc biến độc lập 265.

Bộ lựa chọn chế độ 260 có thể được tạo cấu hình để xác định hoặc lựa chọn phân chia đối với chế độ dự báo khối hiện thời (bao gồm cả không phân chia) và chế độ dự báo (ví dụ chế độ dự báo nội ảnh hoặc liên ảnh) và tạo ra khối dự báo 265 tương ứng, mà được sử dụng cho việc tính toán của khối dư 205 và cho việc tái cấu trúc của khối được tái cấu trúc 215.

Các phương án về bộ lựa chọn chế độ 260 có thể được tạo cấu hình để lựa chọn chế độ phân chia và dự báo (ví dụ từ các chế độ được hỗ trợ bởi hoặc khả dụng đối với bộ lựa chọn chế độ 260), mà cung cấp sự so khớp tốt nhất hoặc nói cách khác phần dư nhỏ nhất (phần dư nhỏ nhất có nghĩa là nén tốt hơn cho việc truyền hoặc lưu trữ), hoặc chi phí báo hiệu nhỏ nhất (chi phí báo hiệu nhỏ nhất có nghĩa là nén tốt hơn cho việc truyền hoặc lưu trữ), hoặc xem xét hoặc cân bằng cả hai. Bộ lựa chọn chế độ 260 có thể được tạo cấu hình để xác định chế độ phân chia và dự báo dựa trên tối ưu hóa biến dạng tỷ lệ (RDO - rate distortion optimization), tức là lựa chọn chế độ dự báo mà cung cấp biến dạng tỷ lệ nhỏ nhất. Các thuật ngữ như “tốt nhất”, “nhỏ nhất”, “tối ưu” vãn vãn trong ngữ cảnh này toàn bộ không nhất thiết phải đề cập đến “tốt nhất”, “nhỏ nhất”, “tối ưu”, vãn vãn mà có thể còn đề cập đến việc hoàn thành hoặc chấm dứt hoặc các tiêu chí lựa chọn như giá trị vượt qua hoặc rơi xuống dưới ngưỡng hoặc các giới hạn khác tiềm tàng việc dẫn đến “lựa chọn dưới mức tối ưu” nhưng giảm độ phức tạp và thời gian xử lý.

Nói cách khác, bộ phân chia 262 có thể được tạo cấu hình để phân chia hình ảnh từ chuỗi video thành chuỗi gồm các đơn vị cây mã hóa (CTU), và CTU 203 có thể còn được phân chia thành các phân chia khối nhỏ hơn hoặc các khối con (mà lại tạo ra các khối), ví dụ lặp đi lặp lại bằng cách sử dụng phân chia cây tứ phân (QT, quad-tree), phân chia cây nhị phân (BT binary tree) hoặc phân chia cây tam phân (TT, triple-tree) hoặc bất kỳ kết hợp nào của chúng, và để thực hiện, ví dụ, dự báo cho mỗi trong số các khối phân chia hoặc các khối con, trong đó chế độ lựa chọn bao gồm lựa chọn cấu trúc cây của khối được phân chia 203 và các chế độ dự báo này được áp dụng cho mỗi trong số các khối phân chia hoặc các khối con.

Trong phần mô tả sau đây sự phân chia (ví dụ bởi bộ phân chia 260) và xử lý dự báo (bởi bộ dự báo nội ảnh 244 và bộ dự báo nội ảnh 254) được thực hiện bởi bộ mã hóa video được lấy làm ví dụ 20 sẽ được giải thích chi tiết hơn.

Phân chia

Bộ phân chia 262 có thể được tạo cấu hình để phân chia hình ảnh từ chuỗi video thành chuỗi gồm các đơn vị cây mã hóa (CTUs), và bộ phân chia 262 có thể phân chia (hoặc chia) đơn vị cây mã hóa (CTU) 203 thành các phân chia nhỏ hơn, ví dụ các khối nhỏ hơn có kích cỡ hình vuông hoặc hình chữ nhật. Đối với hình ảnh mà có ba mảng mẫu, CTU bao gồm khối $N \times N$ có các mẫu độ chói cùng với hai khối tương ứng có các mẫu sắc độ. Kích cỡ cho phép lớn nhất của khối độ chói trong CTU được xác định sẽ là 128×128 khi phát triển mã hóa video đa năng (VVC), nhưng có thể được xác định sẽ là giá trị lớn hơn, ví dụ, 256×256 thay vì 128×128 trong tương lai. Các CTU của hình ảnh có thể được gộp lại/tạo nhóm dưới dạng các lát/các nhóm ô, các ô hoặc các viên. Ô bao lấy vùng hình chữ nhật của hình ảnh, và ô có thể được chia thành một hoặc nhiều viên. Viên bao gồm số lượng các dòng CTU trong ô. Ô mà không được phân chia thành các viên có thể được gọi là viên. Tuy nhiên, viên là tập con của ô và không được gọi là ô. Có hai chế độ gồm các nhóm ô được hỗ trợ trong VVC, tức là chế độ lát/nhóm ô quét mảnh và chế độ lát hình chữ nhật. Trong chế độ nhóm ô quét mảnh, lát/nhóm ô chứa chuỗi các ô khi quét mảnh ô của hình ảnh. Trong chế độ lát hình chữ nhật, lát chứa số lượng các viên của hình ảnh mà nói chung tạo ra vùng hình chữ nhật của hình ảnh. Các viên trong lát hình chữ nhật nằm trong thứ tự quét mảnh viên của lát. Các khối nhỏ hơn (mà có thể còn được gọi là các khối con) có thể còn được phân chia thậm chí thành các phân chia nhỏ hơn. Điều này cũng được đề cập đến là phân chia cây hoặc phân chia cây phân cấp, trong đó khối gốc, ví dụ ở mức

cây gốc 0 (mức phân cấp 0, độ sâu 0), có thể được phân chia đệ quy, ví dụ được phân chia thành hai hoặc nhiều khối có mức cây thấp hơn tiếp theo, ví dụ các nút ở mức cây 1 (mức phân cấp 1, độ sâu 1), trong đó các khối có thể lại được phân chia thành hai hoặc nhiều khối có mức thấp hơn tiếp theo, ví dụ mức cây 2 (mức phân cấp 2, độ sâu 2), vân vân cho đến khi chấm dứt phân chia, ví dụ do tiêu chuẩn chấm dứt được hoàn thành, ví dụ đạt được độ sâu lớn nhất hoặc kích cỡ khối nhỏ nhất. Các khối mà không được phân chia thêm còn được gọi là các khối lá hoặc các nút lá của cây. Cây sử dụng sự phân chia thành hai vùng được gọi là cây nhị phân (BT), cây sử dụng sự phân chia thành ba vùng được gọi là cây tam phân (TT), và cây sử dụng sự phân chia thành bốn vùng được gọi là cây tứ phân (QT).

Ví dụ, đơn vị cây mã hóa (CTU) có thể hoặc bao gồm CTB gồm các mẫu độ chói, hai CTB tương ứng gồm các mẫu sắc độ của hình ảnh mà có ba mảng mẫu, hoặc CTB gồm các mẫu của hình ảnh đơn sắc hoặc hình ảnh mà được lập mã bằng cách sử dụng ba phẳng màu riêng rẽ và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu. Một cách tương ứng, khối cây mã hóa (CTB) có thể là khối $N \times N$ gồm các mẫu dùng cho một vài giá trị N sao cho việc chia thành phần thành các CTB là sự phân chia. Đơn vị mã hóa (CU) có thể hoặc bao gồm khối mã hóa gồm các mẫu độ chói, hai khối mã hóa tương ứng các mẫu sắc độ của hình ảnh mà có ba mảng mẫu, hoặc khối mã hóa có các mẫu của hình ảnh đơn sắc hoặc hình ảnh mà được lập mã bằng cách sử dụng ba phẳng màu riêng rẽ và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu. Một cách tương ứng khối mã hóa (CB) có thể là khối $M \times N$ có các mẫu dùng cho một vài giá trị M và N sao cho việc CTB thành các khối mã hóa là sự phân chia.

Trong các phương án, ví dụ, theo HEVC, đơn vị cây mã hóa (CTU) có thể được chia thành các CU bằng cách sử dụng cấu trúc cây tứ phân được xác định là cây mã hóa. Việc quyết định liệu có mã hóa vùng hình ảnh bằng cách sử dụng dự báo liên ảnh (thời gian) hoặc nội ảnh (không gian) hay không được thực hiện ở mức CU lá. Mỗi CU lá có thể còn được chia thành một, hai hoặc bốn PU theo loại chia PU. Bên trong một PU, cùng một xử lý dự báo được áp dụng và thông tin được truyền đến bộ giải mã trên cơ sở PU. Sau khi thu khối dư bằng cách áp dụng xử lý dự báo dựa trên loại chia PU, CU lá có thể được phân chia thành các đơn vị biến đổi (TU) theo cấu trúc cây tứ phân khác tương tự với cây mã hóa dùng cho CU.

Trong các phương án, ví dụ, theo tiêu chuẩn mã hóa video mới nhất hiện thời đang được phát triển, mà được gọi là mã hóa video đa năng (VVC), cây đa loại kết hợp lồng với cây tứ phân bằng cách sử dụng cấu trúc phân đoạn phân chia nhị phân và tam phân ví dụ được sử dụng để phân chia đơn vị cây mã hóa. Trong cấu trúc cây mã hóa nằm trong đơn vị cây mã hóa, CU có thể có hoặc dạng hình vuông hoặc hình chữ nhật. Ví dụ, đơn vị cây mã hóa (CTU) thứ nhất được phân chia bởi cây tứ phân. Sau đó các nút lá cây tứ phân này có thể còn được phân chia bởi cấu trúc cây đa loại. Có bốn loại phân chia trong cấu trúc cây đa loại, phân chia nhị phân dọc (SPLIT_BT_VER), phân chia nhị phân ngang (SPLIT_BT_HOR), phân chia tam phân dọc (SPLIT_TT_VER), và phân chia tam phân ngang (SPLIT_TT_HOR). Các nút lá của cây đa loại được gọi là các đơn vị mã hóa (CUs), và trừ khi CU quá lớn đối với độ dài biến đổi lớn nhất, sự phân đoạn này được sử dụng để dự báo và xử lý biến đổi mà không cần thêm bất kỳ sự phân chia nào. Điều này có nghĩa là, trong hầu hết các trường hợp, CU, PU và TU có cùng một kích cỡ khối trong cây tứ phân có cấu trúc khối mã hóa cây đa lớp được lồng. Ngoại lệ xảy ra khi độ dài biến đổi được hỗ trợ lớn nhất nhỏ hơn độ rộng hoặc độ cao của thành phần màu của CU. VVC phát triển cơ chế báo hiệu duy nhất cho thông tin phân chia vùng trong cây tứ phân có cấu trúc cây mã hóa đa loại được lồng. Trong cơ chế báo hiệu này, đơn vị cây mã hóa (CTU) được coi là gốc của cây tứ phân và đầu tiên được phân chia bởi cấu trúc cây tứ phân. Mỗi nút lá của cây tứ phân (khi đủ lớn để cho phép) sau đó còn được phân chia bởi cấu trúc cây đa loại. Trong cấu trúc cây đa loại này, cờ thứ nhất (`mtt_split_cu_flag`) được báo hiệu để chỉ báo liệu nút này có còn được phân chia nữa hay không; khi nút còn được phân chia, cờ thứ hai (`mtt_split_cu_vertical_flag`) được báo hiệu để chỉ báo hướng phân chia, và sau đó cờ thứ ba (`mtt_split_cu_binary_flag`) được báo hiệu để chỉ báo liệu sự phân chia này có phải là phân chia nhị phân hoặc phân chia tam phân hay không. Dựa trên các giá trị `mtt_split_cu_vertical_flag` và `mtt_split_cu_binary_flag`, chế độ phân chia cây đa loại (`MttSplitMode`) của CU có thể được dẫn ra bởi bộ giải mã dựa trên quy tắc hoặc bảng xác định trước. Cần lưu ý rằng, đối với thiết kế nhất định, ví dụ, khối độ chói 64×64 và thiết kế đường liên hợp sắc độ 32×32 trong các bộ giải mã phần cứng VVC, phân chia TT còn bị cấm khi hoặc độ rộng hoặc độ cao của khối mã hóa độ chói lớn hơn 64, như được thể hiện trên FIG.6. Phân chia TT bị cấm khi hoặc độ rộng hoặc độ cao của khối mã hóa sắc độ lớn hơn 32. Thiết kế đường liên hợp sắc độ sẽ chia hình ảnh thành các đơn vị dữ liệu đường liên hợp ảo (VPDUs - Virtual pipeline data units) mà được xác định dưới dạng các đơn vị không chồng lên nhau trong hình ảnh. Trong

các bộ giải mã phần cứng, các VPDU kế tiếp được xử lý đồng thời bởi nhiều giai đoạn đường liên hợp. Kích cỡ VPDU nói chung tỷ lệ với kích cỡ bộ đệm trong hầu hết các giai đoạn đường liên hợp, do đó quan trọng là giữ cho kích cỡ VPDU nhỏ. Trong hầu hết các bộ giải mã phần cứng, kích cỡ VPDU có thể được thiết đặt ở kích cỡ khối biến đổi (TB) lớn nhất. Tuy nhiên, trong VVC, phân chia cây tam phân (TT) và cây nhị phân (BT) có thể dẫn đến sự gia tăng kích cỡ các VPDU. Ngoài ra, cần lưu ý rằng, khi một phần của khối nút cây vượt quá đáy hoặc biên hình ảnh bên phải, khối nút cây này bị buộc phải được phân chia cho đến khi tất cả các mẫu của mỗi CU được mã hóa đều nằm ở bên trong các biên hình ảnh.

Ví dụ, công cụ nội phân chia con (ISP, Intra Sub-Partition) có thể chia các khối dự báo nội ảnh độ chói theo chiều dọc hoặc theo chiều ngang thành 2 hoặc 4 vùng con tùy thuộc vào kích cỡ khối.

Trong một ví dụ, bộ lựa chọn chế độ 260 của bộ mã hóa video 20 có thể được tạo cấu hình để thực hiện kết hợp bất kỳ kỹ thuật phân chia nào được mô tả ở đây.

Như được mô tả ở trên, bộ mã hóa video 20 được tạo cấu hình để xác định hoặc lựa chọn chế độ dự báo tốt nhất hoặc tối ưu từ tập các chế độ dự báo (ví dụ được xác định trước). Tập các chế độ dự báo có thể bao gồm, ví dụ, các chế độ dự báo nội ảnh và/hoặc các chế độ dự báo liên ảnh.

Dự báo nội ảnh

Tập các chế độ dự báo nội ảnh có thể bao gồm 35 chế độ dự báo nội ảnh khác nhau, ví dụ các chế độ vô hướng như chế độ DC (hoặc trung bình) và chế độ phẳng, hoặc các chế độ có hướng, ví dụ như được xác định trong HEVC, hoặc thể bao gồm 67 chế độ dự báo nội ảnh khác nhau, ví dụ các chế độ vô hướng như chế độ DC (hoặc trung bình) và chế độ phẳng, hoặc các chế độ có hướng, ví dụ như được xác định đối với VVC. Ví dụ, một vài chế độ dự báo nội ảnh có góc thông thường được thay thế một cách tương thích bằng các chế độ dự báo trong ảnh góc rộng đối với các khối không phải hình vuông, ví dụ như được xác định trong VVC. Ví dụ khác, để tránh các phép chia đối với dự báo DC, chỉ cạnh dài hơn được sử dụng để tính giá trị trung bình đối với các khối không phải hình vuông. Và, các kết quả dự báo nội ảnh của chế độ phẳng có thể còn được cải biến bằng phương pháp kết hợp dự báo nội ảnh phụ thuộc vào vị trí (PDPC - position dependent intra prediction combination).

Bộ dự báo nội ảnh 254 được tạo cấu hình để sử dụng các mẫu được tái cấu trúc của các khối lân cận của cùng một hình ảnh hiện thời để tạo ra khối dự báo nội ảnh 265 theo chế độ dự báo nội ảnh của tập các chế độ dự báo nội ảnh.

Bộ dự báo nội ảnh 254 (hoặc nói chung là bộ lựa chọn chế độ 260) còn được tạo cấu hình để xuất ra các tham số dự báo nội ảnh (hoặc nói chung là chỉ báo thông tin của chế độ dự báo nội ảnh được lựa chọn cho khối) đến bộ mã hóa entropy 270 dưới dạng các thành phần cú pháp 266 để đưa vào dữ liệu hình ảnh được mã hóa 21, sao cho, ví dụ, bộ giải mã video 30 có thể thu nhận và sử dụng các tham số dự báo này cho việc giải mã.

Dự báo liên ảnh (bao gồm dự báo liên lớp)

Tập (hoặc có thể) các chế độ dự báo liên ảnh phụ thuộc vào các hình ảnh tham chiếu khả dụng (tức là ít nhất một phần các hình ảnh được giải mã trước đó, ví dụ được lưu trữ trong DBP 230) và các tham số dự báo liên ảnh khác, ví dụ liệu toàn bộ hình ảnh tham chiếu hoặc chỉ một phần, ví dụ vùng cửa sổ tìm kiếm xung quanh vùng của khối hiện thời, của hình ảnh tham chiếu có được sử dụng để tìm kiếm khối tham chiếu so khớp tốt nhất hay không, và/hoặc ví dụ liệu nội suy điểm ảnh có được áp dụng, ví dụ nội suy một nửa/một phần hai điểm ảnh, một phần tư điểm ảnh và/hoặc 1/16 điểm ảnh, hay không.

Ngoài các chế độ dự báo ở trên ra, chế độ bỏ qua, chế độ trực tiếp và/hoặc chế độ dự báo liên ảnh khác cũng có thể được áp dụng.

Ví dụ, dự báo sáp nhập mở rộng, danh sách ứng viên sáp nhập của chế độ như vậy được cấu trúc bằng cách bao gồm năm loại ứng viên sau đây theo thứ tự: MVP không gian từ các CU lân cận không gian, MVP thời gian từ các CU đồng vị, MVP dựa trên lịch sử từ bảng FIFO, MVP trung bình theo cặp và các MV 0. Và làm mịn vector chuyển động phía bộ giải mã dựa trên so khớp hai chiều (DMVR, bilateral-matching based decoder side motion vector refinement) có thể được áp dụng để gia tăng độ chính xác của các MV của chế độ sáp nhập. Chế độ sáp nhập với MVD (MMVD), mà đến từ chế độ sáp nhập có các chênh lệch vector chuyển động. Cờ MMVD được báo hiệu ngay sau khi gửi cờ bỏ qua và cờ sáp nhập để xác định liệu chế độ MMVD có được sử dụng cho CU hay không. Và sơ đồ độ phân giải tương thích mức CU (AMVR, CU-level adaptive motion vector resolution) có thể được áp dụng. AMVR cho phép MVD của CU sẽ được mã hóa với độ chính xác khác nhau. Tùy thuộc vào chế độ dự báo dùng cho CU hiện thời, các MVD của CU hiện thời có thể được lựa chọn một cách tương thích. Khi CU được mã hóa theo chế độ sáp nhập, chế độ dự báo liên ảnh/nội ảnh kết hợp (CIIP - combined inter/intra prediction) có thể được áp dụng cho CU hiện thời. Lấy trung bình có trọng số của các tín hiệu dự báo liên ảnh và dự báo nội ảnh được thực hiện để thu dự báo CIIP. Dự báo bù chuyển động affine, trường chuyển động

afin của khối được mô tả bằng thông tin chuyển động gồm các vector chuyển động hai điểm điều khiển (4-tham số) hoặc các vector chuyển động ba điểm điều khiển (6-tham số). Dự báo vector chuyển động thời gian dựa trên khối con (SbTMVP, Subblock-based temporal motion vector prediction), mà tương tự với dự báo vector chuyển động thời gian (TMVP, temporal motion vector prediction) trong HEVC, nhưng dự báo các vector chuyển động của các CU con trong CU hiện thời. Dòng quang học hai chiều (BDOF, Bi-directional optical flow), trước đó được gọi là BIO, là phiên bản đơn giản hơn mà yêu cầu ít tính toán hơn nhiều, đặc biệt là về số lượng các phép nhân và kích cỡ của số nhân. Chế độ phân chia tam giác, trong chế độ như vậy, CU được phân chia được phân chia đều thành hai vùng dạng hình tam giác, bằng cách sử dụng hoặc phân chia theo đường chéo hoặc phân chia chống đường chéo. Ngoài ra, chế độ dự báo kép được mở rộng vượt quá việc lấy trung bình đơn giản để cho phép lấy trung bình có trọng số của hai báo hiệu dự báo.

Bộ dự báo liên ảnh 244 có thể bao gồm bộ ước lượng chuyển động (motion estimation - ME) và bộ bù chuyển động (MC - motion compensation) (cả hai đều không được thể hiện trên Fig.2). Bộ ước lượng chuyển động có thể được tạo cấu hình để thu nhận hoặc thu khối hình ảnh 203 (khối hình ảnh hiện thời 203 của hình ảnh hiện thời 17) và hình ảnh được giải mã 231, hoặc ít nhất một hoặc nhiều khối được tái cấu trúc trước đó, ví dụ các khối được tái cấu trúc gồm một hoặc nhiều hình ảnh được giải mã trước đó 231 khác/khác nhau, để ước lượng chuyển động. Ví dụ chuỗi video có thể bao gồm hình ảnh hiện thời và các hình ảnh được giải mã trước đó 231, hoặc nói cách khác, hình ảnh hiện thời và các hình ảnh được giải mã trước đó 231 có thể là một phần của hoặc tạo ra chuỗi các hình ảnh tạo ra chuỗi video.

Bộ mã hóa 20 có thể, ví dụ, được tạo cấu hình để lựa chọn khối tham chiếu từ nhiều khối tham chiếu của cùng một hoặc các hình ảnh khác nhau trong số nhiều hình ảnh khác và cung cấp hình ảnh tham chiếu (hoặc chỉ số hình ảnh tham chiếu) và/hoặc độ dịch (độ dịch không gian) giữa vị trí (các tọa độ x, y) của khối tham chiếu và vị trí của khối hiện thời dưới dạng các tham số dự báo liên ảnh đến bộ ước lượng chuyển động. Độ dịch này còn được gọi là vector chuyển động (MV - motion vector).

Bộ bù chuyển động được tạo cấu hình để thu, ví dụ thu nhận, tham số dự báo liên ảnh và thực hiện dự báo liên ảnh dựa trên hoặc bằng cách sử dụng tham số dự báo liên ảnh để thu khối dự báo liên ảnh 265. Bù chuyển động, được thực hiện bởi bộ

bù chuyển động, có thể liên quan đến tìm nạp hoặc tạo ra khối dự báo dựa trên vector chuyển động/khối được xác định bởi ước lượng chuyển động, có thể thực hiện các nội suy trên độ chính xác điểm ảnh con. Lọc nội suy có thể tạo ra các mẫu điểm ảnh bổ sung từ các mẫu điểm ảnh đã biết, do đó gia tăng tiềm tàng các khối dự báo ứng viên mà có thể được sử dụng để mã hóa khối hình ảnh. Sau khi thu nhận vector chuyển động cho PU của khối hình ảnh hiện thời, bộ bù chuyển động có thể xác định vị trí khối dự báo mà vector chuyển động hướng đến trong một trong số các danh sách hình ảnh tham chiếu.

Bộ bù chuyển động có thể còn tạo ra các thành phần cú pháp liên kết với các khối và các lát video để bộ giải mã video 30 sử dụng khi giải mã các khối hình ảnh của lát video. Ngoài ra hoặc thay thế cho các lát và các thành phần cú pháp tương ứng, các nhóm ô và/hoặc các ô và các thành phần cú pháp tương ứng có thể được tạo ra hoặc được sử dụng.

Mã hóa entropi

Bộ mã hóa entropi 270 được tạo cấu hình để áp dụng, ví dụ, thuật toán hoặc sơ đồ mã hóa entropi (ví dụ sơ đồ mã hóa độ dài thay đổi (VLC, variable length coding), sơ đồ VLC tương thích với ngữ cảnh (CAVLC, context adaptive VLC scheme), sơ đồ mã hóa số học, nhị phân hóa, mã hóa số học nhị phân tương thích với ngữ cảnh (CABAC - context adaptive binary arithmetic coding), mã hóa số học nhị phân tương thích với ngữ cảnh dựa trên cú pháp (SBAC, syntax-based context-adaptive binary arithmetic coding), mã hóa entropi phân chia theo khoảng xác suất (PIPE - probability interval partitioning entropy) hoặc phương pháp hoặc kỹ thuật mã hóa entropi khác) hoặc bỏ qua (không nén) trên các hệ số lượng tử hóa 209, các tham số dự báo liên ảnh, các tham số dự báo nội ảnh, các tham số bộ lọc vòng và/hoặc các thành phần cú pháp khác để thu dữ liệu hình ảnh được mã hóa 21 mà có thể được xuất ra qua đầu ra 272, ví dụ dưới dạng dòng bit được mã hóa 21, sao cho, ví dụ, bộ giải mã video 30 có thể thu nhận và sử dụng các tham số dùng cho việc giải mã. Dòng bit được mã hóa 21 có thể được truyền đến bộ giải mã video 30, hoặc được lưu trữ trong bộ để truyền sau hoặc truy lục bởi bộ giải mã video 30.

Các thay đổi trúc khác của bộ mã hóa video 20 có thể được sử dụng để mã hóa dòng video. Ví dụ, bộ mã hóa dựa trên không biến đổi 20 có thể lượng tử hóa tín hiệu dư một cách trực tiếp mà không cần bộ xử lý biến đổi 206 cho các khối hoặc các khung nhất định. Theo cách thức thực hiện khác, bộ mã hóa 20 có thể có bộ lượng tử hóa 208 và bộ lượng tử hóa ngược 210 được kết hợp thành một bộ.

Bộ giải mã và phương pháp giải mã

Fig.3 thể hiện ví dụ về bộ giải mã video 30 mà được tạo cấu hình để thực hiện các kỹ thuật của sáng chế này. Bộ giải mã video 30 được tạo cấu hình để thu nhận dữ liệu hình ảnh được mã hóa 21 (ví dụ dòng bit được mã hóa 21), ví dụ được mã hóa bởi bộ mã hóa 20, để thu hình ảnh được giải mã 331. Dữ liệu hình ảnh được mã hóa hoặc dòng bit bao gồm thông tin để giải mã dữ liệu hình ảnh được mã hóa, ví dụ dữ liệu mà thể hiện các khối hình ảnh của lát video được mã hóa (và/hoặc các nhóm ô hoặc các ô) và các thành phần cú pháp kết hợp.

Trong ví dụ của FIG.3, bộ giải mã 30 bao gồm bộ giải mã entropi 304, bộ lượng tử hóa ngược 310, bộ xử lý biến đổi ngược 312, bộ tái cấu trúc 314 (ví dụ máy cộng 314), bộ lọc vòng 320, bộ đệm hình ảnh được giải mã (DBP - decoded picture buffer) 330, bộ áp dụng chế độ 360, bộ dự báo liên ảnh 344 và bộ dự báo nội ảnh 354. Bộ dự báo liên ảnh 344 có thể hoặc bao gồm bộ bù chuyển động. Bộ giải mã video 30 có thể, trong một vài ví dụ, thực hiện chuyển tác giải mã thông thường tương hỗ với chuyển tác mã hóa được mô tả đối với bộ mã hóa video 100 từ FIG.2.

Như được giải thích liên quan đến bộ mã hóa 20, bộ lượng tử hóa ngược 210, bộ xử lý biến đổi ngược 212, bộ tái cấu trúc 214, bộ lọc vòng 220, bộ đệm hình ảnh được giải mã (DPB) 230, bộ dự báo liên ảnh 344 và bộ dự báo nội ảnh 354 còn được gọi là tạo ra “bộ giải mã tích hợp” của bộ mã hóa video 20. Do đó, bộ lượng tử hóa ngược 310 về chức năng có thể giống với bộ lượng tử hóa ngược 110, bộ xử lý biến đổi ngược 312 về chức năng có thể giống với bộ xử lý biến đổi ngược 212, bộ tái cấu trúc 314 về chức năng có thể giống với bộ tái cấu trúc 214, bộ lọc vòng 320 về chức năng có thể giống với bộ lọc vòng 220, và bộ đệm hình ảnh được giải mã 330 về chức năng có thể giống với bộ đệm hình ảnh được giải mã 230. Do đó, các cách giải thích mà được đề xuất cho các bộ và các chức năng tương ứng của bộ mã hóa video 20 áp dụng một cách tương ứng cho các bộ và các chức năng tương ứng của bộ giải mã video 30.

Giải mã entropi

Bộ giải mã entropi 304 được tạo cấu hình để phân tích dòng bit 21 (hoặc nói chung là dữ liệu hình ảnh được mã hóa 21) và thực hiện, ví dụ, giải mã entropi đối với dữ liệu hình ảnh được mã hóa 21 để thu, ví dụ, các hệ số lượng tử hóa 309 và/hoặc các tham số mã hóa được giải mã (không được thể hiện trên Fig.3), ví dụ bất kỳ hoặc tất cả các tham số dự báo liên ảnh (ví dụ chỉ số hình ảnh tham chiếu và vector chuyển động), tham số dự báo nội ảnh (ví dụ chế độ dự báo nội ảnh hoặc chỉ số), các tham số biến đổi, các tham số lượng tử hóa, các tham số bộ lọc vòng, và/hoặc

các thành phần cú pháp khác. Bộ giải mã entropi 304 có thể được tạo cấu hình để áp dụng các thuật toán giải mã hoặc các sơ đồ tương ứng với các sơ đồ mã hóa như được mô tả liên quan đến bộ mã hóa entropi 270 của bộ mã hóa 20. Bộ giải mã entropi 304 có thể còn được tạo cấu hình để cung cấp các tham số dự báo liên ảnh, tham số dự báo nội ảnh và/hoặc các thành phần cú pháp khác cho bộ áp dụng chế độ 360 và các tham số khác cho các bộ khác của bộ giải mã 30. Bộ giải mã video 30 có thể thu nhận các thành phần cú pháp ở mức lát video và/hoặc mức khối video. Ngoài ra hoặc thay thế cho các lát và các thành phần cú pháp tương ứng, các nhóm ô và/hoặc các ô và các thành phần cú pháp tương ứng có thể được thu nhận và/hoặc được sử dụng.

Lượng tử hóa ngược

Bộ lượng tử hóa ngược 310 có thể được tạo cấu hình để thu nhận các tham số lượng tử hóa (QP) (hoặc nói chung là thông tin liên quan đến lượng tử hóa ngược) và các hệ số lượng tử hóa từ dữ liệu hình ảnh được mã hóa 21 (ví dụ bằng cách phân tích cú pháp và/hoặc giải mã, ví dụ bằng bộ giải mã entropi 304) và áp dụng dựa trên các tham số lượng tử hóa lượng tử hóa ngược trên các hệ số lượng tử hóa được giải mã 309 để thu các hệ số giải lượng tử hóa 311, mà có thể còn được gọi là các hệ số biến đổi 311. Xử lý lượng tử hóa ngược có thể bao gồm sử dụng tham số lượng tử hóa được xác định bởi bộ mã hóa video 20 cho mỗi khối video trong lát video (hoặc ô hoặc nhóm ô) để xác định mức lượng tử hóa và, tương tự vậy, mức lượng tử hóa ngược mà sẽ được áp dụng.

Biến đổi ngược

Bộ xử lý biến đổi ngược 312 có thể được tạo cấu hình để thu nhận các hệ số giải lượng tử hóa 311, còn được gọi là các hệ số biến đổi 311, và áp dụng biến đổi cho các hệ số giải lượng tử hóa 311 theo thứ tự để thu các khối dư được tái cấu trúc 213 trong cùng một miền. Các khối dư được tái cấu trúc 213 có thể còn được gọi là các khối biến đổi 313. Biến đổi có thể là biến đổi ngược, ví dụ, DCT ngược, DST ngược, biến đổi số nguyên ngược, hoặc xử lý biến đổi ngược tương tự về mặt khái niệm. Bộ xử lý biến đổi ngược 312 có thể còn được tạo cấu hình để thu nhận các tham số biến đổi hoặc thông tin tương ứng từ dữ liệu hình ảnh được mã hóa 21 (ví dụ bằng cách phân tích cú pháp và/hoặc giải mã, ví dụ bằng bộ giải mã entropi 304) để xác định biến đổi sẽ được áp dụng cho các hệ số giải lượng tử hóa 311.

Tái cấu trúc

Bộ tái cấu trúc 314 (ví dụ bộ cộng hoặc máy cộng 314) có thể được tạo cấu hình để bổ sung khối dư được tái cấu trúc 313, vào khối dự báo 365 để thu khối được tái cấu trúc 315 trong cùng một miền, ví dụ bằng cách bổ sung các giá trị mẫu của khối dư được tái cấu trúc 313 và các giá trị mẫu của khối dự báo 365.

Lọc

Bộ lọc vòng 320 (hoặc trong vòng mã hóa hoặc sau vòng mã hóa) được tạo cấu hình để lọc khối được tái cấu trúc 315 để thu khối được lọc 321, ví dụ làm nhẵn các chuyển tiếp điểm ảnh, hoặc nếu không thì cải thiện chất lượng video. Bộ lọc vòng 320 có thể bao gồm một hoặc nhiều bộ lọc vòng chẳng hạn như bộ lọc giải khối, bộ lọc độ dịch tương thích mẫu (SAO) hoặc một hoặc nhiều bộ lọc khác, ví dụ bộ lọc vòng tương thích (ALF), bộ lọc triệt tiêu nhiễu (NSF), hoặc bất kỳ kết hợp nào của chúng. Trong một ví dụ, bộ lọc vòng 220 có thể bao gồm bộ lọc giải khối, bộ lọc SAO và bộ lọc ALF. Thứ tự xử lý lọc có thể là lọc giải khối, SAO và ALF. Trong một ví dụ khác, xử lý được gọi là ánh xạ độ chói có chia tỷ lệ sắc độ (LMCS) (tức là, bộ tái tạo hình dạng trong vòng tương thích) được bổ sung. Xử lý này được thực hiện trước khi giải khối. Trong một ví dụ khác, bộ lọc giải khối xử lý có thể còn được áp dụng cho các mép của khối con bên trong, ví dụ các mép của các khối con afin, các mép của các khối con ATMVP, các mép biến đổi khối con (SBT, sub-block transform) và các mép nội phân chia con (ISP). Mặc dù bộ lọc vòng 320 được thể hiện trên Fig.3 dưới dạng là bộ lọc trong vòng, nhưng trong các cấu hình khác, bộ lọc vòng 320 có thể được thực hiện dưới dạng bộ lọc vòng sau.

Bộ đệm hình ảnh được giải mã

Các khối video được giải mã 321 của hình ảnh sau đó được lưu trữ trong bộ đệm hình ảnh được giải mã 330, mà lưu trữ các hình ảnh được giải mã 331 dưới dạng các hình ảnh tham chiếu dùng để bù chuyển động kế tiếp cho các hình ảnh khác và/hoặc lần lượt xuất ra màn hình.

Bộ giải mã 30 được tạo cấu hình để xuất ra hình ảnh được giải mã 311, ví dụ qua đầu ra 312, để thể hiện hoặc cho người dùng quan sát.

Dự báo

Bộ dự báo liên ảnh 344 có thể giống với bộ dự báo liên ảnh 244 (cụ thể là với bộ bù chuyển động) và bộ dự báo nội ảnh 354 về chức năng có thể giống với bộ dự báo liên ảnh 254, và thực hiện các quyết định phân hoặc phân chia và dự báo dựa trên các tham số phân chia và/hoặc dự báo hoặc thông tin tương ứng được thu nhận từ dữ liệu hình ảnh được mã hóa 21 (ví dụ bằng cách phân tích cú pháp

và/hoặc giải mã, ví dụ bằng bộ giải mã entropi 304). Bộ áp dụng chế độ 360 có thể được tạo cấu hình để thực hiện dự báo (dự báo nội ảnh hoặc liên ảnh (có thể bao gồm dự báo liên lớp)) trên mỗi khối dựa trên được tái cấu trúc các hình ảnh, các khối hoặc các mẫu tương ứng (được lọc hoặc chưa được lọc) để thu khối dự báo 365.

Khi lát video được mã hóa dưới dạng lát được nội mã hóa (I), bộ dự báo nội ảnh 354 của bộ áp dụng chế độ 360 được tạo cấu hình để tạo ra khối dự báo 365 cho khối hình ảnh của lát video hiện thời dựa trên chế độ dự báo nội ảnh được báo hiệu và dữ liệu từ các khối được giải mã trước đó của hình ảnh hiện thời. Khi video hình ảnh được mã hóa dưới dạng lát được liên mã hóa (tức là, B, hoặc P), bộ dự báo liên ảnh 344 (ví dụ bộ bù chuyển động) của bộ áp dụng chế độ 360 được tạo cấu hình để tạo ra các khối dự báo 365 cho khối video của lát video hiện thời dựa trên các vector chuyển động và các thành phần cú pháp khác được thu nhận từ bộ giải mã entropi 304. Đối với dự báo liên ảnh, các khối dự báo có thể được tạo ra từ một trong số các hình ảnh tham chiếu trong một trong số các danh sách hình ảnh tham chiếu. Bộ giải mã video 30 có thể cấu trúc các danh sách khung tham chiếu, Danh sách 0 và Danh sách 1, bằng cách sử dụng các kỹ thuật cấu trúc mặc định dựa trên các hình ảnh tham chiếu được lưu trữ trong DPB 330. Điều giống hoặc tương tự có thể được áp dụng cho hoặc bởi các phương án bằng cách sử dụng các nhóm ô (ví dụ các nhóm ô video) và/hoặc các ô (ví dụ các ô video) ngoài ra hoặc thay cho các lát (ví dụ các lát video) ra, ví dụ video có thể được mã hóa bằng cách sử dụng các nhóm ô I, P hoặc B và /hoặc các ô.

Bộ áp dụng chế độ 360 được tạo cấu hình để xác định thông tin dự báo cho khối video của lát video hiện thời bằng cách phân tích cú pháp các vector chuyển động hoặc thông tin liên quan và các thành phần cú pháp khác, và sử dụng thông tin dự báo để tạo ra các khối dự báo cho khối video hiện thời đang được giải mã. Ví dụ, bộ áp dụng chế độ 360 sử dụng một vài trong số các thành phần cú pháp được thu nhận để xác định chế độ dự báo (ví dụ, dự báo nội ảnh hoặc liên ảnh) được sử dụng để mã hóa các khối video của lát video, loại dự báo liên ảnh lát (ví dụ, lát B, lát P, hoặc lát GPB), thông tin cấu trúc cho một hoặc nhiều danh sách hình ảnh tham chiếu đối với lát, các vector chuyển động cho mỗi khối video được liên mã hóa của lát, trạng thái dự báo liên ảnh cho mỗi khối video được liên mã hóa của lát, và thông tin khác để giải mã các khối video trong lát video hiện thời. Điều giống hoặc tương tự có thể được áp dụng cho hoặc bởi các phương án bằng cách sử dụng các nhóm ô (ví dụ các nhóm ô video) và/hoặc các ô (ví dụ các ô video)

ngoài ra hoặc thay cho các lát (ví dụ các lát video) ra, ví dụ video có thể được mã hóa bằng cách sử dụng các nhóm ô I, P hoặc B và/hoặc các ô.

Các phương án về bộ giải mã video 30 như được thể hiện trên FIG.3 có thể được tạo cấu hình để phân chia và/hoặc giải mã hình ảnh bằng cách sử dụng các lát (còn được gọi là các lát video), trong đó hình ảnh có thể được phân chia thành hoặc được giải mã bằng cách sử dụng một hoặc nhiều lát (thông thường không chồng lên nhau), và mỗi lát có thể bao gồm một hoặc nhiều khối (ví dụ các CTU) hoặc một hoặc nhiều nhóm gồm các khối (ví dụ các ô (H.265/HEVC và VVC) hoặc các viên (VVC)).

Các phương án về bộ giải mã video 30 như được thể hiện trên FIG.3 có thể được tạo cấu hình để phân chia và/hoặc giải mã hình ảnh bằng cách sử dụng các lát/các nhóm ô (còn được gọi là các nhóm ô video) và/hoặc các ô (còn được gọi là các ô video), trong đó hình ảnh có thể được phân chia hoặc được giải mã bằng cách sử dụng một hoặc nhiều lát/các nhóm ô (thông thường không chồng lên nhau), và mỗi lát/nhóm ô có thể bao gồm, ví dụ một hoặc nhiều khối (ví dụ các CTU) hoặc một hoặc nhiều ô, trong đó mỗi ô, ví dụ có thể có dạng hình chữ nhật và có thể bao gồm một hoặc nhiều khối (ví dụ các CTU), ví dụ các khối hoàn chỉnh hoặc phân đoạn.

Các thay đổi khác của bộ giải mã video 30 có thể được sử dụng để giải mã dữ liệu hình ảnh được mã hóa 21. Ví dụ, bộ giải mã 30 có thể tạo ra đầu ra dòng video mà không cần bộ lọc vòng 320. Ví dụ, bộ giải mã dựa trên không biến đổi 30 có thể lượng tử hóa ngược tín hiệu dư một cách trực tiếp mà không cần bộ xử lý biến đổi ngược 312 đối với các khối hoặc các khung nhất định. Theo cách thức thực hiện khác, bộ giải mã video 30 có thể có bộ lượng tử hóa ngược 310 và bộ xử lý biến đổi ngược 312 được kết hợp thành một bộ.

Cần hiểu rằng, trong bộ mã hóa 20 và bộ giải mã 30, kết quả xử lý của bước hiện tại có thể còn được xử lý và sau đó được xuất ra đến bước kế tiếp. Ví dụ, sau khi lọc nội suy, việc dẫn ra vectơ chuyển động hoặc lọc vòng, thao tác khác, chẳng hạn như tách hoặc dịch chuyển, có thể được thực hiện trên kết quả xử lý của việc lọc nội suy, việc dẫn ra vectơ chuyển động hoặc lọc vòng.

Cần lưu ý rằng thao tác khác có thể được áp dụng cho các vectơ chuyển động được dẫn ra của khối hiện thời (bao gồm nhưng không giới hạn ở điểm điều khiển các vectơ chuyển động của chế độ afin, các vectơ chuyển động khối con trong afin, các chế độ phẳng, ATMVP, các vectơ chuyển động thời gian, và vân vân). Ví dụ, giá trị của vectơ chuyển động bị giới hạn ở khoảng xác định trước theo bit

đại diện của nó. Nếu bit đại diện của vector chuyển động là bitDepth, thì sau đó khoảng này là $-2^{(\text{bitDepth}-1)} \sim 2^{(\text{bitDepth}-1)-1}$, trong đó “^” có nghĩa là lũy thừa. Ví dụ, nếu bitDepth được thiết đặt bằng 16, thì khoảng này là $-32768 \sim 32767$; nếu bitDepth được thiết đặt bằng 18, thì khoảng này là $-131072 \sim 131071$. Ví dụ, giá trị của vector chuyển động được dẫn ra (ví dụ các MV trong số bốn khối con 4x4 trong một khối 8x8) bị giới hạn sao cho chênh lệch lớn nhất giữa các phần nguyên của bốn khối con 4x4 không lớn hơn N các điểm ảnh, chẳng hạn như không lớn hơn 1 điểm ảnh. Ở đây sáng chế đề xuất hai phương pháp giới hạn vector chuyển động theo bitDepth.

FIG.4 là sơ đồ giản lược của thiết bị mã hóa video 400 theo một phương án của sáng chế. Thiết bị mã hóa video 400 là phù hợp để thực hiện các phương án được bộc lộ như được mô tả ở đây. Trong một phương án, thiết bị mã hóa video 400 có thể là bộ giải mã chẳng hạn như bộ giải mã video 30 của FIG.1A hoặc bộ mã hóa chẳng hạn như bộ mã hóa video 20 của FIG.1A.

Thiết bị mã hóa video 400 bao gồm các cổng vào 410 (hoặc các cổng đầu vào 410) và các bộ thu nhận (Rx) 420 dùng để thu nhận dữ liệu; bộ xử lý, bộ logic, hoặc bộ xử lý trung tâm (CPU) 430 để xử lý dữ liệu; các bộ truyền (Tx) 440 và các cổng ra 450 (hoặc các cổng đầu ra 450) dùng để truyền dữ liệu; và bộ nhớ 460 dùng để lưu trữ dữ liệu. Thiết bị mã hóa video 400 có thể còn bao gồm các thành phần quang-điện (optical-to-electrical, OE) và các thành phần điện-quang (electrical-to-optical, EO) được ghép nối với các cổng vào 410, các bộ thu nhận 420, các bộ truyền 440, và các cổng ra 450 dùng để đưa ra hoặc đưa vào báo hiệu quang học hoặc điện.

Bộ xử lý 430 được thực hiện bởi phần cứng và phần mềm. Bộ xử lý 430 có thể được thực hiện dưới dạng một hoặc nhiều vi mạch CPU, các lõi (ví dụ, như bộ xử lý đa lõi), các FPGA, các ASIC, và các DSP. Bộ xử lý 430 truyền thông vào các cổng vào 410, các bộ thu nhận 420, các bộ truyền 440, các cổng ra 450, và bộ nhớ 460. Bộ xử lý 430 bao gồm môđun mã hóa 470. Môđun mã hóa 470 thực hiện các phương án bộc lộ được mô tả ở trên. Ví dụ, môđun mã hóa 470 thực hiện, xử lý, chuẩn bị, hoặc cung cấp các hoạt động mã hóa khác nhau. Việc đưa vào môđun mã hóa 470 do đó cung cấp sự cải tiến đáng kể chức năng của thiết bị mã hóa video 400 và tạo hiệu quả biến đổi của thiết bị mã hóa video 400 sang trạng thái khác. Ngoài ra, môđun mã hóa 470 được thực hiện dưới dạng các lệnh được lưu trữ trong bộ nhớ 460 và được thực hiện bởi bộ xử lý 430.

Bộ nhớ 460 có thể bao gồm một hoặc nhiều đĩa, ổ băng, và các đĩa dạng rắn và có thể được sử dụng làm thiết bị lưu trữ dữ liệu quá dòng, để lưu trữ các chương trình khi các chương trình như vậy được lựa chọn để thực hiện, và lưu trữ các lệnh và dữ liệu mà được đọc khi thực hiện chương trình. Bộ nhớ 460 có thể là, ví dụ, khả biến và/hoặc bất biến và có thể bộ nhớ chỉ đọc (ROM), bộ nhớ truy nhập ngẫu nhiên (RAM), bộ nhớ định địa chỉ nội dung tam phân (TCAM, ternary content-addressable memory), và/hoặc bộ nhớ truy nhập ngẫu nhiên tĩnh (SRAM - static random-access memory).

Fig.5 là sơ đồ khối giản lược của thiết bị 500 mà có thể được sử dụng như một trong hai thiết bị nguồn 12 và thiết bị đích 14 từ Fig.1 theo một phương án được lấy làm ví dụ.

Bộ xử lý 502 trong thiết bị 500 có thể là bộ xử lý trung tâm. Ngoài ra, bộ xử lý 502 có thể là bất kỳ loại thiết bị nào khác, hoặc các thiết bị, có khả năng thao tác hoặc xử lý thông tin hiện có hoặc sau này khi được phát triển. Mặc dù các cách thức thực hiện được bộc lộ có thể được thực hiện với một bộ xử lý như được thể hiện, ví dụ, nhưng bộ xử lý 502, có thể mà có ưu điểm về hiệu quả và tốc độ thu được bằng cách sử dụng nhiều hơn một bộ xử lý.

Bộ nhớ 504 trong thiết bị 500 có thể là thiết bị bộ nhớ chỉ đọc (ROM) hoặc thiết bị bộ nhớ truy nhập ngẫu nhiên (RAM) trong một cách thức thực hiện. Bất kỳ loại thiết bị lưu trữ nào khác cũng có thể được sử dụng làm bộ nhớ 504. Bộ nhớ 504 có bao gồm mã và dữ liệu 506 mà được truy nhập bởi bộ xử lý 502 bằng cách sử dụng kênh truyền 512. Bộ nhớ 504 ngoài ra có thể bao gồm hệ điều hành 508 và các chương trình ứng dụng 510, các chương trình ứng dụng 510 bao gồm ít nhất một chương trình mà cho phép bộ xử lý 502 thực hiện các phương pháp được mô tả ở đây. Ví dụ, các chương trình ứng dụng 510 có thể bao gồm các ứng dụng 1 đến N, mà ngoài ra bao gồm ứng dụng mã hóa video mà thực hiện các phương pháp được mô tả ở đây.

Thiết bị 500 có thể còn bao gồm một hoặc nhiều thiết bị đầu ra, chẳng hạn như màn hình 518. Màn hình 518 có thể, trong một ví dụ, là màn hình nhạy chạm mà kết hợp màn hình với chi tiết nhạy chạm mà có thể hoạt động để cảm biến các đầu vào chạm. Màn hình 518 có thể được ghép nối với bộ xử lý 502 qua kênh truyền 512.

Mặc dù được mô tả ở đây dưới dạng một kênh truyền, nhưng kênh truyền 512 của thiết bị 500 có thể bao gồm nhiều kênh truyền. Ngoài ra, bộ lưu trữ thứ hai 514 có thể được ghép nối một cách trực tiếp với các thành phần khác của thiết bị 500

hoặc có thể được truy nhập qua mạng và có thể bao gồm một bộ tích hợp chẳng hạn như thẻ nhớ hoặc nhiều bộ phận chẳng hạn như nhiều thẻ nhớ. Thiết bị 500 do đó có thể được thực hiện theo một loạt các cấu hình.

Mã hóa mở rộng

Mã hóa mở rộng bao gồm mở rộng chất lượng (mở rộng PSNR), mở rộng không gian, các mở rộng khác. Ví dụ, như được thể hiện trên Fig.6, chuỗi có thể được lấy mẫu xuống đến phiên bản độ phân giải không gian thấp. Cả phiên bản độ phân giải không gian thấp và phiên bản độ phân giải không gian gốc (độ phân giải không gian cao) đều sẽ được mã hóa. Và thông thường, độ phân giải không gian thấp sẽ được mã hóa trước tiên, và sẽ được sử dụng để tham chiếu cho độ phân giải không gian cao được mã hóa sau.

Để mô tả thông tin của các lớp (số lượng, sự phụ thuộc, xuất ra), có VPS (Tập tham số video) được xác định như sau:

video_parameter_set_rbsp() {	Ký hiệu mô tả
...	
vps_max_layers_minus1h	u(6)
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(j = 0; j < i; j++)	
vps_direct_dependency_flag[i][j]	u(1)
}	
}	
...	

}	
---	--

vps_max_layers_minus1 plus 1 xác định số lượng các lớp được cho phép lớn nhất trong mỗi CVS tham chiếu đến VPS.

vps_all_independent_layers_flag bằng 1 xác định rằng tất cả các lớp trong CVS đều được mã hóa một cách độc lập mà không sử dụng dự báo liên lớp. **vps_all_independent_layers_flag** bằng 0 xác định rằng một hoặc nhiều lớp trong CVS có thể sử dụng dự báo liên lớp. Khi không có mặt, giá trị **vps_all_independent_layers_flag** được suy ra sẽ bằng 1. Khi **vps_all_independent_layers_flag** bằng 1, giá trị **vps_independent_layer_flag[i]** được suy ra sẽ bằng 1. Khi **vps_all_independent_layers_flag** bằng 0, giá trị **vps_independent_layer_flag[0]** được suy ra sẽ bằng 1.

vps_layer_id[i] xác định giá trị **nuh_layer_id** của lớp thứ *i*. Đối với bất kỳ hai giá trị nguyên không âm là *m* và *n*, khi *m* nhỏ hơn *n*, giá trị **vps_layer_id[m]** sẽ 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 liên ảnh. **vps_independent_layer_flag[i]** bằng 0 xác định rằng lớp có chỉ số *i* có thể sử dụng dự báo liên ảnh và **vps_layer_dependency_flag[i]** có mặt trong VPS.

vps_direct_dependency_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_dependency_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_dependency_flag[i][j] không có mặt đối với *i* và *j* nằm trong khoảng từ 0 đến **vps_max_layers_minus1**, bao gồm cả hai giá trị này, suy ra nó sẽ bằng 0.

Biến **DirectDependentLayerIdx[i][j]**, xác định lớp phụ thuộc trực tiếp thứ *j* của lớp thứ *i*, được dẫn ra như sau:

```

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

```

(7-2)

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

```
for( i = 0; i <= vps_max_layers_minus1; i++ )      (7-3)
```

```
    GeneralLayerIdx[ vps_layer_id[ i ] ] = i
```

phần mô tả đơn giản như sau:

vps_max_layers_minus1 add 1 có nghĩa là số lượng của các lớp.

vps_all_independent_layers_flag chỉ báo liệu tất cả các lớp có được mã hóa một cách độc lập hay không.

vps_layer_id[i] chỉ báo ID lớp của lớp thứ i.

vps_independent_layer_flag[i] chỉ báo liệu lớp thứ i có được mã hóa một cách độc lập hay không.

vps_direct_dependency_flag[i][j] chỉ báo liệu lớp thứ j có được sử dụng để tham chiếu đối với lớp thứ i hay không.

Trong đó các thành phần cú pháp vps_independent_layer_flag[i] và vps_direct_dependency_flag[i][j] đều là thông tin dự báo liên lớp của các lớp, trong đó i, và j là ký hiệu nhận dạng lớp, các lớp khác nhau tương ứng với ký hiệu nhận dạng lớp khác nhau.

Quản lý DPB và chú thích hình ảnh tham chiếu

Để quản lý các hình ảnh tham chiếu đó khi xử lý giải mã, các hình ảnh được giải mã cần phải được giữ trong bộ đệm giải mã hình ảnh (DPB), để sử dụng tham chiếu đối với hình ảnh giải mã sau đó. Để chỉ báo các hình ảnh đó, thông tin số đếm thứ tự hình ảnh (POC - picture order count) của chúng cần phải báo hiệu trong đoạn đầu lát một cách trực tiếp hoặc một cách gián tiếp. Thông thường, có danh sách hình ảnh tham chiếu, list0 và list1. Và, chỉ số hình ảnh tham chiếu cũng cần phải được bao gồm để báo hiệu hình ảnh trong danh sách này. Đối với dự báo đơn nhất, các hình ảnh tham chiếu được tìm nạp từ một danh sách hình ảnh tham chiếu, đối với dự báo kép, các hình ảnh tham chiếu được tìm nạp từ hai danh sách hình ảnh tham chiếu.

Tất cả các hình ảnh tham chiếu đều được lưu trữ trong DPB. Tất cả các hình ảnh trong DPB được chú thích là “được sử dụng cho tham chiếu dài hạn”, “được sử dụng cho tham chiếu ngắn hạn”, hoặc “không được sử dụng để tham chiếu”, và chỉ một trong ba trạng thái này. Sau khi hình ảnh được chú thích là “không được sử dụng để tham chiếu”, nó sẽ không được sử dụng để tham chiếu nữa. Nếu cũng không cần lưu trữ để xuất ra, thì sau đó nó có thể được loại bỏ khỏi DPB. Trạng thái của các hình ảnh tham chiếu có thể được báo hiệu trong đoạn đầu lát, hoặc có thể được dẫn ra từ thông tin đoạn đầu lát.

Phương pháp quản lý hình ảnh tham chiếu mới được đề xuất, được gọi là phương pháp RPL (danh sách hình ảnh tham chiếu - reference picture list). RPL sẽ đề xuất toàn bộ tập hoặc các tập hình ảnh tham chiếu cho hình ảnh mã hóa hiện thời, hình ảnh tham chiếu trong tập hình ảnh tham chiếu được sử dụng cho hình ảnh hiện thời hoặc hình ảnh giải mã tương lai (sau, hoặc sau đó). Do đó, RPL phản ánh các thông tin các hình ảnh trong DPB, ngay cả hình ảnh tham chiếu không được sử dụng để tham chiếu đối với hình ảnh hiện thời, nếu sẽ được sử dụng để tham chiếu đối với hình ảnh sau đó, thì cần lưu trữ trong RPL.

Sau khi hình ảnh được tái cấu trúc, sẽ được lưu trữ trong DPB, và được chú thích là “được sử dụng cho tham chiếu ngắn hạn” một cách mặc định. Hoạt động quản lý DPB sẽ bắt đầu sau khi phân tích cú pháp thông tin RPL trong đoạn đầu lát.

Cấu trúc danh sách hình ảnh tham chiếu

Thông tin hình ảnh tham chiếu có thể được báo hiệu qua đoạn đầu lát. Ngoài ra, có thể có một vài ứng viên RPL trong tập tham số chuỗi (SPS), trong trường hợp này, đoạn đầu lát có thể bao gồm chỉ số RPL để thu được thông tin RPL cần thiết, mà không cần báo hiệu toàn bộ cấu trúc cú pháp. Hoặc, toàn bộ cấu trúc cú pháp RPL có thể được báo hiệu trong đoạn đầu lát.

Giới thiệu phương pháp RPL

Để tiết kiệm các bit chi phí của báo hiệu RPL, có thể một vài ứng viên RPL trong SPS. Hình ảnh có thể sử dụng chỉ số RPL (`ref_pic_list_idx[ i ]`) để thu được thông tin RPL của nó từ SPS. Các ứng viên RPL được báo hiệu như sau:

<code>seq_parameter_set_rbsp() {</code>	Ký hiệu mô tả
<code>...</code>	
<code>rpl1_same_as_rpl0_flag</code>	<code>u(1)</code>
<code>for(i = 0; i < !rpl1_same_as_rpl0_flag ? 2 : 1; i++) {</code>	
<code>num_ref_pic_lists_in_sps[i]</code>	<code>ue(v)</code>
<code>for(j = 0; j < num_ref_pic_lists_in_sps[i]; j++)</code>	
<code>ref_pic_list_struct(i, j)</code>	
<code>}</code>	
<code>...</code>	

Các ngữ nghĩa là như sau:

rpl1_same_as_rpl0_flag bằng 1 xác định rằng các cấu trúc cú pháp **num_ref_pic_lists_in_sps[1]** và **ref_pic_list_struct(1, rplsIdx)** không có mặt và sau đây áp dụng:

- Giá trị **num_ref_pic_lists_in_sps[1]** được suy ra sẽ bằng với giá trị **num_ref_pic_lists_in_sps[0]**.
- Giá trị của mỗi trong số các thành phần cú pháp trong **ref_pic_list_struct(1, rplsIdx)** được suy ra sẽ bằng với giá trị của thành phần cú pháp tương ứng trong **ref_pic_list_struct(0, rplsIdx)** đối với **rplsIdx** nằm trong khoảng từ 0 đến **num_ref_pic_lists_in_sps[0] – 1**.

num_ref_pic_lists_in_sps[i] xác định số lượng các cấu trúc cú pháp **ref_pic_list_struct(listIdx, rplsIdx)** có **listIdx** bằng **i** được bao gồm trong SPS. Giá trị **num_ref_pic_lists_in_sps[i]** sẽ nằm trong khoảng từ 0 đến 64, bao gồm cả hai giá trị này.

Bên cạnh việc thu được thông tin RPL dựa trên chỉ số RPL từ SPS, thông tin RPL có thể được báo hiệu trong đoạn đầu lát.

slice_header() {	Ký hiệu mô tả
...	
if((nal_unit_type != IDR_W_RADL && nal_unit_type != IDR_N_LP) sps_idr_rpl_present_flag) {	
for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && !pps_ref_pic_list_sps_idc[i] && (i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 && (i == 0 (i == 1 && rpl1_idx_present_flag)))	

ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i])	
for(j = 0; j < NumLtrpEntries[i][RplsIdx[i]]; j++) {	
if(ltrp_in_slice_header_flag[i][RplsIdx[i]])	
slice_poc_lsb_lt[i][j]	u(v)
delta_poc_msb_present_flag[i][j]	u(1)
if(delta_poc_msb_present_flag[i][j])	
delta_poc_msb_cycle_lt[i][j]	ue(v)
}	
}	
...	

ref_pic_list_sps_flag[i] bằng 1 xác định rằng danh sách hình ảnh tham chiếu i của lát hiện thời được dẫn ra dựa trên một trong số các cấu trúc cú pháp

ref_pic_list_struct(listIdx, rplsIdx) có listIdx bằng i trong SPS.

ref_pic_list_sps_flag[i] bằng 0 xác định rằng danh sách hình ảnh tham chiếu i của lát hiện thời được dẫn ra dựa trên cấu trúc cú pháp

ref_pic_list_struct(listIdx, rplsIdx) có listIdx bằng i mà một cách trực tiếp được bao gồm trong các đoạn đầu lát của hình ảnh hiện thời.

Khi **ref_pic_list_sps_flag[i]** không có mặt, sau đây áp dụng:

- Nếu **num_ref_pic_lists_in_sps[i]** bằng 0, thì giá trị **ref_pic_list_sps_flag[i]** được suy ra sẽ bằng 0.
- Nếu không thì (**num_ref_pic_lists_in_sps[i]** lớn hơn 0), nếu **rpl1_idx_present_flag** bằng 0, thì giá trị **ref_pic_list_sps_flag[1]** được suy ra sẽ bằng với **ref_pic_list_sps_flag[0]**.
- Nếu không thì, giá trị **ref_pic_list_sps_flag[i]** được suy ra sẽ bằng với **pps_ref_pic_list_sps_idx[i] - 1**.

ref_pic_list_idx[i] xác định chỉ số, trong danh sách các cấu trúc cú pháp

ref_pic_list_struct(listIdx, rplsIdx) có listIdx bằng i được bao gồm trong SPS,

danh sách cấu trúc cú pháp **ref_pic_list_struct(listIdx, rplsIdx)** có listIdx bằng i

mà được sử dụng để dẫn ra danh sách hình ảnh tham chiếu i của hình ảnh hiện thời. Thành phần cú pháp $\text{ref_pic_list_idx}[i]$ được thể hiện bởi các bit $\text{Ceil}(\text{Log2}(\text{num_ref_pic_lists_in_sps}[i]))$. Khi không có mặt, giá trị $\text{ref_pic_list_idx}[i]$ được suy ra sẽ bằng 0. Giá trị $\text{ref_pic_list_idx}[i]$ sẽ nằm trong khoảng từ 0 đến $\text{num_ref_pic_lists_in_sps}[i] - 1$, bao gồm cả hai giá trị này. Khi $\text{ref_pic_list_sps_flag}[i]$ bằng 1 và $\text{num_ref_pic_lists_in_sps}[i]$ bằng 1, giá trị $\text{ref_pic_list_idx}[i]$ được suy ra sẽ bằng 0. Khi $\text{ref_pic_list_sps_flag}[i]$ bằng 1 và $\text{rpl1_idx_present_flag}$ bằng 0, giá trị $\text{ref_pic_list_idx}[1]$ được suy ra sẽ bằng với $\text{ref_pic_list_idx}[0]$.

Biến $\text{RplsIdx}[i]$ được dẫn ra như sau:

$$\begin{aligned} \text{RplsIdx}[i] &= \text{ref_pic_list_sps_flag}[i] ? \text{ref_pic_list_idx}[i] : \\ &\text{num_ref_pic_lists_in_sps}[i] \quad (7-95) \end{aligned}$$

slice_poc_lsb_lt $[i][j]$ xác định giá trị của mô đun số đếm thứ tự hình ảnh MaxPicOrderCntLsb của mục nhập LTRP thứ j trong danh sách hình ảnh tham chiếu thứ i . Độ dài của thành phần cú pháp $\text{slice_poc_lsb_lt}[i][j]$ là $\log_2 \text{max_pic_order_cnt_lsb_minus4} + 4$ bits.

Biến $\text{PocLsbLt}[i][j]$ được dẫn ra như sau:

$$\begin{aligned} \text{PocLsbLt}[i][j] &= \text{ltrp_in_slice_header_flag}[i][\text{RplsIdx}[i]] ? \quad (7-96) \\ &\text{slice_poc_lsb_lt}[i][j] : \\ &\text{rpls_poc_lsb_lt}[\text{listIdx}][\text{RplsIdx}[i]][j] \end{aligned}$$

delta_poc_msb_present_flag $[i][j]$ bằng 1 xác định rằng $\text{delta_poc_msb_cycle_lt}[i][j]$ có mặt. $\text{delta_poc_msb_present_flag}[i][j]$ bằng 0 xác định rằng $\text{delta_poc_msb_cycle_lt}[i][j]$ không có mặt.

Hãy coi prevTid0Pic là hình ảnh trước đó theo thứ tự giải mã mà có nuh_layer_id giống như hình ảnh hiện thời, có TemporalId bằng 0, và không phải là hình ảnh RAL hoặc RADL. Hãy coi setOfPrevPocVals là tập bao gồm:

- PicOrderCntVal của prevTid0Pic ,
- PicOrderCntVal của mỗi hình ảnh mà được tham chiếu đến bởi các mục nhập trong $\text{RefPicList}[0]$ hoặc $\text{RefPicList}[1]$ của prevTid0Pic và có nuh_layer_id giống với hình ảnh hiện thời,
- PicOrderCntVal của mỗi hình ảnh mà theo sau prevTid0Pic theo thứ tự giải mã, có nuh_layer_id giống với hình ảnh hiện thời, và đứng trước hình ảnh hiện thời theo thứ tự giải mã.

Khi có nhiều hơn một giá trị trong setOfPrevPocVals mà mô đun giá trị

MaxPicOrderCntLsb bằng với PocLsbLt[i][j], giá trị delta_poc_msb_present_flag[i][j] sẽ bằng 1.

delta_poc_msb_cycle_lt[i][j] xác định giá trị của biến FullPocLt[i][j] như sau:

```

if( j == 0 )
    DeltaPocMsbCycleLt[ i ][ j ] = delta_poc_msb_cycle_lt[ i ][ j ]
else
    DeltaPocMsbCycleLt[ i ][ j ] = delta_poc_msb_cycle_lt[ i ][ j ] +
    DeltaPocMsbCycleLt[ i ][ j - 1 ]
FullPocLt[ i ][ j ] = PicOrderCntVal - DeltaPocMsbCycleLt[ i ][ j ] *
MaxPicOrderCntLsb -
    ( PicOrderCntVal & ( MaxPicOrderCntLsb - 1 ) ) +
    PocLsbLt[ i ][ j ]

```

Giá trị delta_poc_msb_cycle_lt[i][j] sẽ nằm trong khoảng từ 0 đến $2^{(32 - \log_2 \text{max_pic_order_cnt_lsb_minus4} - 4)}$, bao gồm cả hai giá trị này. Khi không có mặt, giá trị delta_poc_msb_cycle_lt[i][j] được suy ra sẽ bằng 0.

Cấu trúc cú pháp của RPL như sau:

ref_pic_list_struct (listIdx, rplsIdx) {	Ký hiệu mô tả
num_ref_entries [listIdx][rplsIdx]	ue(v)
if(long_term_ref_pics_flag)	
ltrp_in_slice_header_flag [listIdx][rplsIdx]	u(1)
for(i = 0, j = 0; i < num_ref_entries[listIdx][rplsIdx]; i++) {	
if(inter_layer_ref_pics_present_flag)	
inter_layer_ref_pic_flag [listIdx][rplsIdx][i]	u(1)
if(!inter_layer_ref_pics_flag[listIdx][rplsIdx][i]) {	
if(long_term_ref_pics_flag)	
st_ref_pic_flag [listIdx][rplsIdx][i]	u(1)
if(st_ref_pic_flag[listIdx][rplsIdx][i]) {	
abs_delta_poc_st [listIdx][rplsIdx][i]	ue(v)
if(AbsDeltaPocSt[listIdx][rplsIdx][i] > 0)	
strp_entry_sign_flag [listIdx][rplsIdx][i]	u(1)
} else if(!ltrp_in_slice_header_flag[listIdx][rplsIdx])	
rpls_poc_lsb_lt [listIdx][rplsIdx][j++]	u(v)
} else	
ilrp_idc [listIdx][rplsIdx][i]	ue(v)
}	
}	

num_ref_entries[listIdx][rplsIdx] xác định số lượng các mục nhập trong cấu trúc cú pháp **ref_pic_list_struct**(listIdx, rplsIdx). Giá trị **num_ref_entries**[listIdx][rplsIdx] sẽ nằm trong khoảng từ 0 đến **sps_max_dec_pic_buffering_minus1** + 14, bao gồm cả hai giá trị này.

ltrp_in_slice_header_flag[listIdx][rplsIdx] bằng 0 xác định rằng các POC LSB của các mục nhập LTRP trong cấu trúc cú pháp

`ref_pic_list_struct( listIdx, rplsIdx )` có mặt trong cấu trúc cú pháp

`ref_pic_list_struct( listIdx, rplsIdx )`.

`ltrp_in_slice_header_flag[ listIdx ][ rplsIdx ]` bằng 1 xác định rằng các POC LSB của các mục nhập LTRP trong cấu trúc cú pháp

`ref_pic_list_struct( listIdx, rplsIdx )` không có mặt trong cấu trúc cú pháp

`ref_pic_list_struct( listIdx, rplsIdx )`.

`inter_layer_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]` bằng 1 xác định rằng mục nhập thứ *i* trong cấu trúc cú pháp `ref_pic_list_struct( listIdx, rplsIdx )` là mục nhập ILRP. **`inter_layer_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]`** bằng 0 xác định rằng mục nhập thứ *i* trong cấu trúc cú pháp `ref_pic_list_struct( listIdx, rplsIdx )` không phải là mục nhập ILRP. Khi không có mặt, giá trị

`inter_layer_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]` được suy ra sẽ bằng 0.

`st_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]` bằng 1 xác định rằng mục nhập thứ *i* trong cấu trúc cú pháp `ref_pic_list_struct( listIdx, rplsIdx )` là mục nhập STRP.

`st_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]` bằng 0 xác định rằng mục nhập thứ *i* trong cấu trúc cú pháp `ref_pic_list_struct( listIdx, rplsIdx )` là mục nhập LTRP.

Khi `inter_layer_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]` bằng 0 và `st_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]` không có mặt, giá trị của `st_ref_pic_flag[ listIdx ][ rplsIdx ][ i ]` được suy ra sẽ bằng 1.

Biến `NumLtrpEntries[ listIdx ][ rplsIdx ]` được dẫn ra như sau:

```
for( i = 0, NumLtrpEntries[ listIdx ][ rplsIdx ] = 0; i <
  num_ref_entries[ listIdx ][ rplsIdx ]; i++ )
  if( !inter_layer_ref_pic_flag[ listIdx ][ rplsIdx ][ i ] &&
    !st_ref_pic_flag[ listIdx ][ rplsIdx ][ i ] )
    NumLtrpEntries[ listIdx ][ rplsIdx ]++
```

(7-120)

`abs_delta_poc_st[ listIdx ][ rplsIdx ][ i ]` xác định giá trị của biến

`AbsDeltaPocSt[ listIdx ][ rplsIdx ][ i ]` như sau:

```
if( sps_weighted_pred_flag || sps_weighted_bipred_flag )
  AbsDeltaPocSt[ listIdx ][ rplsIdx ][ i ] =
    abs_delta_poc_st[ listIdx ][ rplsIdx ][ i ] (7-121)
else
  AbsDeltaPocSt[ listIdx ][ rplsIdx ][ i ] =
    abs_delta_poc_st[ listIdx ][ rplsIdx ][ i ] + 1
```

Giá trị `abs_delta_poc_st[ listIdx ][ rplsIdx ][ i ]` sẽ nằm trong khoảng từ 0 đến $2^{15} - 1$, bao gồm cả hai giá trị này.

strp_entry_sign_flag[listIdx][rplsIdx][i] bằng 1 xác định rằng mục nhập thứ i trong cấu trúc cú pháp **ref_pic_list_struct**(listIdx, rplsIdx) có giá trị lớn hơn hoặc bằng 0. **strp_entry_sign_flag**[listIdx][rplsIdx][i] bằng 0 xác định rằng mục nhập thứ i trong cấu trúc cú pháp **ref_pic_list_struct**(listIdx, rplsIdx) có giá trị nhỏ hơn 0. Khi không có mặt, giá trị **strp_entry_sign_flag**[listIdx][rplsIdx][i] được suy ra sẽ bằng 1.

Danh sách **DeltaPocValSt**[listIdx][rplsIdx] được dẫn ra như sau:

```
for( i = 0; i < num_ref_entries[ listIdx ][ rplsIdx ]; i++ )
    if( !inter_layer_ref_pic_flag[ listIdx ][ rplsIdx ][ i ] &&
        st_ref_pic_flag[ listIdx ][ rplsIdx ][ i ] )
        DeltaPocValSt[ listIdx ][ rplsIdx ][ i ] = (
            strp_entry_sign_flag[ listIdx ][ rplsIdx ][ i ] ) ?
            AbsDeltaPocSt[ listIdx ][ rplsIdx ][ i ] : 0 -
            AbsDeltaPocSt[ listIdx ][ rplsIdx ][ i ]
```

(7-122)

rpls_poc_lsb_lt[listIdx][rplsIdx][i] xác định giá trị của mô đun số đếm thứ tự hình ảnh **MaxPicOrderCntLsb** của hình ảnh được tham chiếu đến bởi mục nhập thứ i trong cấu trúc cú pháp **ref_pic_list_struct**(listIdx, rplsIdx). Độ dài của thành phần cú pháp **rpls_poc_lsb_lt**[listIdx][rplsIdx][i] là $\log_2_max_pic_order_cnt_lsb_minus4 + 4$ bits.

Một vài mô tả chung về cấu trúc RPL.

Đối với mỗi danh sách, có cấu trúc RPL. Đầu tiên,

num_ref_entries[listIdx][rplsIdx] được báo hiệu để chỉ báo số lượng của các hình ảnh tham chiếu trong danh sách.

ltrp_in_slice_header_flag[listIdx][rplsIdx] được sử dụng để chỉ báo liệu thông tin LSB (Least Significant Bit) có được báo hiệu trong đoạn đầu lát hay không.

Nếu hình ảnh hiện thời tham chiếu không phải là hình ảnh tham chiếu liên lớp, thì sau đó **st_ref_pic_flag**[listIdx][rplsIdx][i] chỉ báo liệu nó có phải là hình ảnh tham chiếu dài hạn hay không. Nếu là hình ảnh tham chiếu ngắn hạn, thì sau đó thông tin POC (**abs_delta_poc_st** và **strp_entry_sign_flag**) được báo hiệu.

Nếu **ltrp_in_slice_header_flag**[listIdx][rplsIdx] bằng 0, thì sau đó **rpls_poc_lsb_lt**[listIdx][rplsIdx][i] được sử dụng để dẫn ra thông tin LSB của hình ảnh hiện thời tham chiếu. MSB (Most Significant Bit – bit có nghĩa nhất) có thể được dẫn ra một cách trực tiếp, hoặc được dẫn ra dựa trên thông tin (**delta_poc_msb_present_flag**[ei][j] và **delta_poc_msb_cycle_lt**[ei][j]) trong đoạn đầu lát.

Xử lý giải mã đối với cấu trúc các danh sách hình ảnh tham chiếu

Xử lý này được gọi ra khi bắt đầu xử lý giải mã đối với mỗi lát của hình ảnh không phải IDR.

Các hình ảnh tham chiếu được định địa chỉ qua các chỉ số tham chiếu. Chỉ số tham chiếu là chỉ số trong danh sách hình ảnh tham chiếu. Khi giải mã lát I, không có danh sách hình ảnh tham chiếu nào được sử dụng khi giải mã dữ liệu lát. Khi giải mã lát P, chỉ danh sách hình ảnh tham chiếu 0 (tức là, RefPicList[0]), được sử dụng khi giải mã dữ liệu lát. Khi giải mã lát B, cả danh sách hình ảnh tham chiếu 0 và danh sách hình ảnh tham chiếu 1 (tức là, RefPicList[1]) được sử dụng khi giải mã dữ liệu lát.

Khi bắt đầu xử lý giải mã đối với mỗi lát của hình ảnh không phải IDR, các danh sách hình ảnh tham chiếu RefPicList[0] và RefPicList[1] được dẫn ra. Các danh sách hình ảnh tham chiếu được sử dụng khi chú thích các hình ảnh tham chiếu như được xác định trong khoản 8.3.3 hoặc khi giải mã dữ liệu lát.

LƯU Ý 1 – Đối với lát I của hình ảnh không phải IDR mà không phải là lát thứ nhất của hình ảnh, RefPicList[0] và RefPicList[1] có thể được dẫn ra nhằm mục đích kiểm tra sự tương thích dòng bit, nhưng sự dẫn ra của chúng là không cần thiết cho việc giải mã hình ảnh hiện thời hoặc các hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã. Đối với lát P lát mà không phải là lát thứ nhất của hình ảnh, RefPicList[1] có thể được dẫn ra nhằm mục đích kiểm tra sự tương thích dòng bit, nhưng việc dẫn ra của nó là không cần thiết cho việc giải mã hình ảnh hiện thời hoặc các hình ảnh sau hình ảnh hiện thời theo thứ tự giải mã.

Các danh sách hình ảnh tham chiếu RefPicList[0] và RefPicList[1] được cấu trúc như sau:

```
for( i = 0; i < 2; i++ ) {
    for( j = 0, k = 0, pocBase = PicOrderCntVal; j <
num_ref_entries[ i ][ RplsIdx[ i ] ]; j++ ) {
        if( !inter_layer_ref_pic_flag[ i ][ RplsIdx[ i ] ][ j ] ) {
            if( st_ref_pic_flag[ i ][ RplsIdx[ i ] ][ j ] ) {
                RefPicPocList[ i ][ j ] =
pocBase - DeltaPocValSt[ i ][ RplsIdx[ i ] ][ j ]
                if( nếu có hình ảnh tham chiếu picA trong DPB có nuh_layer_id
giống như hình ảnh hiện thời
                    và PicOrderCntVal bằng với RefPicPocList[ i ][ j ] )
                    RefPicList[ i ][ j ] = picA
```

```

else
    RefPicList[ i ][ j ] = " không có hình ảnh tham chiếu "    (8-5)
    pocBase = RefPicPocList[ i ][ j ]
} else {
    if( !delta_poc_msb_cycle_lt[ i ][ k ] ) {
        if( có tham chiếu picA trong DPB có nuh_layer_id giống như
hình ảnh hiện thời và
            PicOrderCntVal & ( MaxPicOrderCntLsb - 1 ) bằng
với PocLsbLt[ i ][ k ] )
            RefPicList[ i ][ j ] = picA
        else
            RefPicList[ i ][ j ] = "không có hình ảnh tham chiếu"
            RefPicLtPocList[ i ][ j ] = PocLsbLt[ i ][ k ]
    } else {
        if( có tham chiếu picA trong DPB có nuh_layer_id giống như
hình ảnh hiện thời và
            PicOrderCntVal bằng với FullPocLt[ i ][ k ] )
            RefPicList[ i ][ j ] = picA
        else
            RefPicList[ i ][ j ] = "không có hình ảnh tham chiếu"
            RefPicLtPocList[ i ][ j ] = FullPocLt[ i ][ k ]
    }
    k++
}
} else {
    layerIdx =
DirectDependentLayerIdx[ GeneralLayerIdx[ nuh_layer_id ] ][ ilrp_idc[ i ]
[ RplsIdx ][ j ] ]
    refPicLayerId = vps_layer_id[ layerIdx ]
    if( có hình ảnh tham chiếu picA trong DPB có nuh_layer_id bằng
với refPicLayerId và
        PicOrderCntVal giống như hình ảnh hiện thời )
        RefPicList[ i ][ j ] = picA
    else
        RefPicList[ i ][ j ] = "không có hình ảnh tham chiếu"
}

```

```

    }
}

```

Sau khi các RPL được cấu trúc, trong đó refPicLayerId là ký hiệu nhận dạng lớp của ILRP, PicOrderCntVal là giá trị POC, xử lý chú thích như sau:

Xử lý giải mã dùng để chú thích hình ảnh tham chiếu

Xử lý này được gọi ra một lần trên mỗi hình ảnh, sau khi giải mã đoạn đầu lát và xử lý giải mã dùng cho cấu trúc danh sách hình ảnh tham chiếu đối với lát như được xác định trong khoản 8.3.2, nhưng trước khi giải mã lát dữ liệu. Xử lý này có thể dẫn đến một hoặc nhiều hình ảnh tham chiếu trong DPB được chú thích là "không được sử dụng để tham chiếu" hoặc "được sử dụng cho tham chiếu dài hạn".

Hình ảnh được giải mã trong DPB có thể được chú thích là "không được sử dụng để tham chiếu", "được sử dụng cho tham chiếu ngắn hạn" hoặc "được sử dụng cho tham chiếu dài hạn", nhưng chỉ một trong số ba chú thích này được chú thích tại bất kỳ thời điểm nào cho trước khi hoạt động xử lý giải mã. Việc gán một trong số các chú thích này cho hình ảnh ngụ ý loại bỏ các chú thích khác khi áp dụng. Khi hình ảnh được gọi là được chú thích là "được sử dụng để tham chiếu", nói chung đề cập đến hình ảnh được chú thích là "được sử dụng cho tham chiếu ngắn hạn" hoặc "được sử dụng cho tham chiếu dài hạn" (nhưng không phải là cả hai).

Các STRP và các ILRP được nhận dạng bởi các giá trị nuh_layer_id và PicOrderCntVal của chúng. Các LTRP được nhận dạng bởi giá trị nuh_layer_ids và các $\text{Log}_2(\text{MaxLtPicOrderCntLsb})$ LSB trong số các giá trị PicOrderCntVal của chúng.

Nếu hình ảnh hiện thời là hình ảnh CLVSS, thì tất cả các hình ảnh tham chiếu hiện thời trong DPB (nếu có) đều có nuh_layer_id giống với hình ảnh hiện thời được chú thích là "không được sử dụng để tham chiếu".

Nếu không thì, sau đây áp dụng:

- Đối với mỗi mục nhập LTRP trong RefPicList[0] or RefPicList[1], khi hình ảnh được tham chiếu là STRP có nuh_layer_id giống với hình ảnh hiện thời, hình ảnh này được chú thích là "được sử dụng cho tham chiếu dài hạn".
- Mỗi hình ảnh tham chiếu có nuh_layer_id giống với hình ảnh hiện thời trong DPB mà không được tham chiếu đến bởi bất kỳ mục nhập nào trong RefPicList[0] hoặc RefPicList[1] được chú thích là "không được sử dụng để tham chiếu".

- Đối với mỗi mục nhập ILRP trong RefPicList[0] hoặc RefPicList[1], hình ảnh được tham chiếu được chú thích là "được sử dụng cho tham chiếu dài hạn"

Cần lưu ý ở đây rằng, ILRP (hình ảnh tham chiếu liên lớp) được chú thích là “được sử dụng cho tham chiếu dài hạn”.

Có hai cú pháp trong SPS mà liên quan đến thông tin tham chiếu liên lớp.

seq_parameter_set_rbsp() {	Ký hiệu mô tả
...	
sps_video_parameter_set_id	u(4)
...	
long_term_ref_pics_flag	u(1)
inter_layer_ref_pics_present_flag	u(1)
sps_idr_rpl_present_flag	u(1)
rpl1_same_as_rpl0_flag	u(1)
...	

sps_video_parameter_set_id, khi lớn hơn 0, xác định giá trị **sps_video_parameter_set_id** cho VPS được tham chiếu đến bởi SPS. Khi **sps_video_parameter_set_id** bằng 0, SPS không tham chiếu đến VPS và không có VPS nào được tham chiếu đến khi giải mã mỗi CVS tham chiếu đến SPS.

long_term_ref_pics_flag bằng 0 xác định rằng không có LTRP nào được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong CVS. **long_term_ref_pics_flag** bằng 1 xác định rằng các LTRP có thể được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong CVS.

inter_layer_ref_pics_present_flag bằng 0 xác định rằng không có ILRP nào được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong CVS. **inter_layer_ref_pics_flag** bằng 1 xác định rằng các ILRP có thể được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong CVS. Khi

sps_video_parameter_set_id bằng 0, giá trị inter_layer_ref_pics_present_flag được suy ra sẽ bằng 0.

Phần mô tả đơn giản như dưới đây:

long_term_ref_pics_flag được sử dụng để chỉ báo liệu LTRP có thể được sử dụng khi xử lý giải mã.

inter_layer_ref_pics_present_flag được sử dụng để chỉ báo liệu ILRP có thể được sử dụng khi xử lý giải mã.

Do đó, khi inter_layer_ref_pics_present_flag bằng 1, có thể có ILRP mà được sử dụng khi xử lý giải mã, và được chú thích là "được sử dụng cho tham chiếu dài hạn". Trong trường hợp này, có LTRP được sử dụng khi xử lý giải mã, ngay cả long_term_ref_pics_flag bằng 0. Do đó long_term_ref_pics_flag có mâu thuẫn về ngữ nghĩa.

Trong phương pháp hiện tại, một vài thành phần cú pháp dùng cho thông tin tham chiếu liên lớp luôn được báo hiệu, mà không cần xem xét đến chỉ số của lớp hiện thời. Sáng chế đề xuất bổ sung một vài điều kiện vào các thành phần cú pháp để cải thiện hiệu quả báo hiệu.

Do long_term_ref_pics_flag chỉ được sử dụng để điều khiển việc phân tích cú pháp ltrp_in_slice_header_flag và st_ref_pic_flag, nên ngữ nghĩa được cải biến để điều khiển việc phân tích cú pháp là phân tích cú pháp các cờ trong RPL.

Các thành phần cú pháp dùng cho thông tin tham chiếu liên lớp được báo hiệu có xem xét đến chỉ số của lớp hiện thời. Nếu thông tin có thể được dẫn ra bởi chỉ số của lớp hiện thời, thì thông tin này không thiết phải được báo hiệu.

Do long_term_ref_pics_flag chỉ được sử dụng để điều khiển việc phân tích cú pháp ltrp_in_slice_header_flag và st_ref_pic_flag, nên ngữ nghĩa được cải biến để điều khiển việc phân tích cú pháp là phân tích cú pháp các cờ trong RPL.

Các thành phần cú pháp dùng cho thông tin tham chiếu liên lớp được báo hiệu xem xét đến chỉ số của lớp hiện thời. Nếu thông tin có thể được dẫn ra bởi chỉ số của lớp hiện thời, thì thông tin này không cần thiết phải được báo hiệu.

Phương án thứ nhất (cải biến ngữ nghĩa long_term_ref_pics_flag để loại bỏ mâu thuẫn giữa LTRP và ILRP)

Do long_term_ref_pics_flag chỉ được sử dụng để điều khiển việc phân tích cú pháp ltrp_in_slice_header_flag và st_ref_pic_flag, nên ngữ nghĩa được cải biến như sau:

long_term_ref_pics_flag bằng 1, xác định rằng ltrp_in_slice_header_flag, và

st_ref_pic_flag có mặt trong các cấu trúc cú pháp ref_pic_list_struct(listIdx, rplsIdx). long_term_ref_pics_flag bằng 0 xác định rằng các thành phần cú pháp này không có mặt trong các cấu trúc cú pháp ref_pic_list_struct(listIdx, rplsIdx). bằng 0 xác định rằng không có LTRP nào được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong CVS. long_term_ref_pics_flag bằng 1 xác định rằng Các LTRP có thể được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong CVS.

Ngoài ra, ngữ nghĩa có thể được cải biến để loại trừ ILRP như sau:

long_term_ref_pics_flag bằng 0 xác định rằng không có LTRP nào được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong CVS. long_term_ref_pics_flag bằng 1 xác định rằng các LTRP có thể được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong CVS. Ở đây LTRP này không bao gồm ILRP (hình ảnh tham chiếu liên lớp).

Phương án thứ hai

Lưu ý ở đây rằng, khi i bằng 1, mà có nghĩa là layer1 cần phải tham chiếu đến lớp khác. Trong khi chỉ layer0 có thể là lớp tham chiếu, do đó, vps_direct_dependency_flag[i][j] không cần phải được báo hiệu. Chỉ khi i lớn hơn 1, vps_direct_dependency_flag[i][j] mới cần được được báo hiệu.

video_parameter_set_rbsp() {	Ký hiệu mô tả
vps_video_parameter_set_id	u(4)
vps_max_layers_minus1	u(6)
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(i > 1 && !vps_independent_layer_flag[i])	
for(j = 0; j < i; j++)	

vps_direct_dependency_flag[i][j]	u(1)
}	
}	
if(vps_max_layers_minus1 > 0) {	
...	

vps_direct_dependency_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_dependency_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_dependency_flag[i][j] không có mặt đối với i và j nằm trong khoảng từ 0 đến **vps_max_layers_minus1**, bao gồm cả hai giá trị này, nếu i bằng 1 và **vps_independent_layer_flag[i]** bằng 0, thì

vps_direct_dependency_flag[i][j] được suy ra sẽ bằng 1, nếu không thì, suy ra sẽ bằng 0.

Phương án thứ ba

Lưu ý ở đây rằng nếu **sps_video_parameter_set_id** (SPS-level syntax element) bằng 0, sau đó có nghĩa là không có nhiều lớp, do đó không cần phải báo hiệu **inter_layer_ref_pics_flag** (thành phần cú pháp cho phép liên lớp), và cờ mặc định là 0.

seq_parameter_set_rbsp() {	Ký hiệu mô tả
...	
sps_video_parameter_set_id	u(4)
...	
long_term_ref_pics_flag	u(1)
if(sps_video_parameter_set_id > 0)	
inter_layer_ref_pics_present_flag	u(1)
sps_idr_rpl_present_flag	u(1)
rpl1_same_as_rpl0_flag	u(1)

...	

inter_layer_ref_pics_present_flag bằng 0 xác định rằng không có ILRP nào được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong CVS. **inter_layer_ref_pics_flag** bằng 1 xác định rằng các ILRP có thể được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong CVS. Khi **sps_video_parameter_set_id** bằng 0 **inter_layer_ref_pics_flag** không có mặt, giá trị **inter_layer_ref_pics_present_flag** suy ra sẽ bằng 0.

Lưu ý ở đây rằng, khi **GeneralLayerIdx[nuh_layer_id]** bằng 0, sau đó lớp hiện thời là lớp thứ 0, không thể tham chiếu đến bất kỳ lớp nào khác. Do đó, không cần phải báo hiệu **inter_layer_ref_pics_present_flag**, và giá trị này mặc định là 0.

seq_parameter_set_rbsp() {	Ký hiệu mô tả
...	
sps_video_parameter_set_id	u(4)
...	
long_term_ref_pics_flag	u(1)
if(GeneralLayerIdx[nuh_layer_id] > 0)	
inter_layer_ref_pics_present_flag	u(1)
sps_idr_rpl_present_flag	u(1)
rpl1_same_as_rpl0_flag	u(1)
...	

inter_layer_ref_pics_present_flag bằng 0 xác định rằng không có ILRP nào được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong CVS. **inter_layer_ref_pics_flag** bằng 1 xác định rằng các ILRP có thể được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong CVS. Khi

~~sps_video_parameter_set_id~~ bằng 0 ~~inter_layer_ref_pics_flag~~ không có mặt, giá trị ~~inter_layer_ref_pics_present_flag~~ suy ra sẽ bằng 0.

Khi mã hóa cả hai trường hợp được đề cập ở trên, ví dụ ứng dụng khác được thể hiện dưới đây:

seq_parameter_set_rbsp() {	Ký hiệu mô tả
...	
sps_video_parameter_set_id	u(4)
...	
long_term_ref_pics_flag	u(1)
if(sps_video_parameter_set_id > 0 && GeneralLayerIdx[nuh_layer_id] > 0)	
inter_layer_ref_pics_present_flag	u(1)
sps_idr_rpl_present_flag	u(1)
rpl1_same_as_rpl0_flag	u(1)
...	

inter_layer_ref_pics_present_flag bằng 0 xác định rằng không có ILRP nào được sử dụng cho dự báo liên ảnh của hình ảnh được lập mã bất kỳ trong CVS. **inter_layer_ref_pics_flag** bằng 1 xác định rằng các ILRP có thể được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã trong CVS. Khi ~~sps_video_parameter_set_id~~ bằng 0 ~~inter_layer_ref_pics_flag~~ không có mặt, giá trị ~~inter_layer_ref_pics_present_flag~~ suy ra sẽ bằng 0.

Phương án thứ tư (thông tin tham chiếu liên lớp được báo hiệu xem xét đến chỉ số của lớp hiện thời, để loại bỏ thông tin báo hiệu dư thừa, để cải thiện hiệu quả mã hóa).

Lưu ý ở đây rằng, khi GeneralLayerIdx[nuh_layer_id] bằng 1, sau đó lớp hiện thời là layer1, và nó chỉ có thể tham chiếu đến layer0, trong khi ilrp_idc của layer0 phải là 0. Do đó, không cần phải báo hiệu ilrp_idc trong trường hợp này.

ref_pic_list_struct(listIdx, rplsIdx) {	Ký hiệu mô tả
num_ref_entries[listIdx][rplsIdx]	ue(v)
if(long_term_ref_pics_flag)	
ltrp_in_slice_header_flag[listIdx][rplsIdx]	u(1)
for(i = 0, j = 0; i < num_ref_entries[listIdx][rplsIdx]; i++) {	
if(inter_layer_ref_pics_present_flag)	
inter_layer_ref_pic_flag[listIdx][rplsIdx][i]	u(1)
if(!inter_layer_ref_pics_flag[listIdx][rplsIdx][i]) {	
if(long_term_ref_pics_flag)	
st_ref_pic_flag[listIdx][rplsIdx][i]	u(1)
if(st_ref_pic_flag[listIdx][rplsIdx][i]) {	
abs_delta_poc_st[listIdx][rplsIdx][i]	ue(v)
if(AbsDeltaPocSt[listIdx][rplsIdx][i] > 0)	
strp_entry_sign_flag[listIdx][rplsIdx][i]	u(1)
} else if(!ltrp_in_slice_header_flag[listIdx][rplsIdx])	
rpls_poc_lsb_lt[listIdx][rplsIdx][j++]	u(v)
} else if(GeneralLayerIdx[nuh_layer_id] > 1)	
ilrp_idc[listIdx][rplsIdx][i]	ue(v)
}	
}	

ilrp_idc[listIdx][rplsIdx][i] xác định chỉ số, đối với danh sách các lớp phụ thuộc trực tiếp, ILRP của mục nhập thứ i trong cấu trúc cú pháp **ref_pic_list_struct(listIdx, rplsIdx)** trên danh sách của các lớp phụ thuộc trực tiếp này. Giá trị **ilrp_idc[listIdx][rplsIdx][i]** sẽ nằm trong khoảng từ 0 đến **GeneralLayerIdx[nuh_layer_id] - 1**, bao gồm cả hai giá trị này. Khi **GeneralLayerIdx[nuh_layer_id]** bằng 1, giá trị **ilrp_idc[listIdx][rplsIdx][i]**

được suy ra sẽ bằng 0.

Phương án thứ năm

Lưu ý ở đây rằng một phần hoặc tất cả các phương án phương án 1 ~ phương án 4 có thể được kết hợp để tạo ra phương án mới.

Ví dụ, phương án 1 + phương án 2 + phương án 3 + phương án 4, hoặc phương án 2 + phương án 3 + phương án 4, hoặc các kết hợp khác.

Phần mô tả sau đây giải thích các ứng dụng của phương pháp mã hóa cũng như phương pháp giải mã như được thể hiện trong các phương án được đề cập ở trên, và hệ thống sử dụng các phương pháp này.

FIG.7 là sơ đồ khối thể hiện hệ thống cung cấp nội dung 3100 dùng để nhận diện dịch vụ phân bổ nội dung. Hệ thống cung cấp nội dung 3100 này bao gồm thiết bị chụp 3102, thiết bị đầu cuối 3106, và một cách tùy chọn bao gồm màn hình 3126. Thiết bị chụp 3102 truyền thông với thiết bị đầu cuối 3106 qua liên kết truyền thông 3104. Liên kết truyền thông có thể bao gồm kênh truyền thông 13 được mô tả ở trên. Liên kết truyền thông 3104 bao gồm nhưng không giới hạn ở WIFI, Ethernet, cáp, kết nối không dây (3G/4G/5G), USB, hoặc bất kỳ loại kết hợp nào của chúng, hoặc tương tự.

Thiết bị chụp 3102 tạo ra dữ liệu, và có thể mã hóa dữ liệu bằng phương pháp mã hóa như được thể hiện trong các phương án ở trên. Ngoài ra, thiết bị chụp 3102 có thể phân bổ dữ liệu đến máy chủ phát trực tuyến (không được thể hiện trên các FIG), và máy chủ mã hóa dữ liệu và truyền dữ liệu được mã hóa đến thiết bị đầu cuối 3106. Thiết bị chụp 3102 bao gồm nhưng không giới hạn ở camera, điện thoại thông minh hoặc đầu in, máy tính hoặc máy tính xách tay, hệ thống hội nghị qua video, PDA, thiết bị lắp trên phương tiện, hoặc kết hợp bất kỳ của chúng, hoặc tương tự. Ví dụ, thiết bị chụp 3102 có thể bao gồm thiết bị nguồn 12 như được mô tả ở trên. Khi dữ liệu bao gồm video, bộ mã hóa video 20 được bao gồm trong thiết bị chụp 3102 có thể thực sự thực hiện mã hóa video việc xử lý. Khi dữ liệu bao gồm audio (tức là, giọng nói), bộ mã hóa audio được bao gồm trong thiết bị chụp 3102 thực ra có thể thực hiện xử lý mã hóa audio. Đối với một vài kịch bản thực tế, thiết bị chụp 3102 phân bổ dữ liệu video và audio được mã hóa bằng cách ghép kênh chúng với nhau. Đối với các kịch bản khác, ví dụ trong hệ thống hội nghị qua video, dữ liệu audio được mã hóa và dữ liệu video được mã hóa không được ghép kênh. Thiết bị chụp 3102 phân bổ dữ liệu audio được mã hóa và dữ liệu video được mã hóa đến thiết bị đầu cuối 3106 một cách riêng rẽ.

Trong hệ thống cung cấp nội dung 3100, thiết bị đầu cuối 310 thu nhận và sao chép lại dữ liệu được mã hóa. Thiết bị đầu cuối 3106 có thể là thiết bị có dữ liệu thu nhận và khôi phục khả năng, chẳng hạn như điện thoại thông minh hoặc đầu in 3108, máy tính hoặc máy tính xách tay 3110, thiết bị ghi video mạng (NVR - network video recorder)/ thiết bị ghi video kỹ thuật số (DVR - digital video recorder) 3112, tivi 3114, đầu thu giải mã tín hiệu (set top box - STB) 3116, hệ thống hội nghị qua video 3118, hệ thống giám sát video 3120, thiết bị hỗ trợ số cá nhân (PDA - personal digital assistant) 3122, thiết bị lắp trên phương tiện 3124, hoặc kết hợp bất kỳ của chúng, hoặc tương tự có thể giải mã dữ liệu được mã hóa được đề cập ở trên. Ví dụ, thiết bị đầu cuối 3106 có thể bao gồm thiết bị đích 14 như được mô tả ở trên. Khi dữ liệu được mã hóa bao gồm video, bộ giải mã video 30 được bao gồm trong thiết bị đầu cuối được ưu tiên thực hiện giải mã video. Khi dữ liệu được mã hóa bao gồm audio, bộ giải mã audio được bao gồm trong thiết bị đầu cuối được ưu tiên thực hiện xử lý giải mã audio.

Đối với thiết bị đầu cuối có màn hình, ví dụ, điện thoại thông minh hoặc đầu in 3108, máy tính hoặc máy tính xách tay 3110, thiết bị ghi video mạng (NVR)/ thiết bị ghi video kỹ thuật số (DVR) 3112, Tivi 3114, thiết bị hỗ trợ số cá nhân (PDA) 3122, hoặc thiết bị lắp trên phương tiện 3124, thiết bị đầu cuối có thể nạp dữ liệu được giải mã vào màn hình của chúng. Đối với thiết bị đầu cuối không được trang bị màn hình, ví dụ, STB 3116, hệ thống hội nghị qua video 3118, hoặc hệ thống giám sát video 3120, màn hình bên ngoài 3126 được chạm vào để thu nhận và thể hiện dữ liệu được giải mã.

Khi mỗi thiết bị trong hệ thống này thực hiện mã hóa hoặc giải mã, thiết bị mã hóa hình ảnh hoặc thiết bị giải mã hình ảnh, như được thể hiện trong các phương án được đề cập ở trên, có thể được sử dụng.

FIG.8 là sơ đồ thể hiện cấu trúc của thiết bị đầu cuối 3106 được lấy làm ví dụ. Sau khi thiết bị đầu cuối 3106 thu nhận dòng từ thiết bị chụp 3102, bộ tiến hành giao thức 3202 phân tích giao thức truyền của dòng. Giao thức này bao gồm nhưng không giới hạn ở giao thức phát trực tuyến theo thời gian thực (RTSP, Real Time Streaming Protocol), giao thức truyền tải siêu văn bản (HTTP, Hyper Text Transfer Protocol), giao thức phát trực tuyến trực tiếp HTTP (HLS, HTTP Live streaming protocol), MPEG-DASH, giao thức truyền tải theo thời gian thực (RTP, Real-time Transport protocol), giao thức nhắn tin theo thời gian thực (RTMP, Real Time Messaging Protocol), hoặc kết hợp bất kỳ của chúng, hoặc tương tự.

Sau khi bộ tiến hành giao thức 3202 xử lý dòng, tệp dòng được tạo ra. Tệp này được xuất ra đến bộ tách kênh 3204. Bộ tách kênh 3204 có thể phân tách dữ liệu được ghép kênh thành dữ liệu audio được mã hóa và dữ liệu video được mã hóa. Như được mô tả ở trên, đối với một vài kịch bản thực tế, ví dụ trong hệ thống hội nghị qua video, dữ liệu audio được mã hóa và dữ liệu video được mã hóa không được ghép kênh. Trong tình huống này, dữ liệu được mã hóa được truyền đến bộ giải mã video 3206 và bộ giải mã audio 3208 mà không qua bộ tách kênh 3204.

Qua xử lý tách kênh, dòng sơ cấp (elementary stream - ES) video, ES audio, và một cách tùy chọn phụ đề được tạo ra. Bộ giải mã video 3206, mà bao gồm bộ giải mã video 30 như được giải thích trong các phương án được đề cập ở trên, giải mã ES video bằng phương pháp giải mã như được thể hiện trong các phương án được đề cập ở trên để tạo ra khung video, và nạp dữ liệu này vào bộ đồng bộ 3212. Bộ giải mã audio 3208, giải mã ES audio để tạo ra khung audio, và nạp dữ liệu này vào bộ đồng bộ 3212. Ngoài ra, khung video có thể lưu trữ trong bộ đệm (không được thể hiện trên Fig.Y) trước khi nạp nó vào bộ đồng bộ 3212. Tương tự, khung audio có thể lưu trữ trong bộ đệm (không được thể hiện trên Fig.Y) trước khi nạp nó vào bộ đồng bộ 3212.

Bộ đồng bộ 3212 đồng bộ khung video và khung audio, và cung cấp video/audio đến màn hình video/audio 3214. Ví dụ, bộ đồng bộ 3212 đồng bộ việc trình bày thông tin video và audio. Thông tin có thể mã hóa trong cú pháp bằng cách sử dụng các dấu thời gian liên quan đến việc trình bày dữ liệu audio và trực quan được mã hóa và các dấu thời gian liên quan đến việc phân phối chính dòng dữ liệu.

Nếu phụ đề được bao gồm trong dòng, thì bộ giải mã phụ đề 3210 giải mã phụ đề, và đồng bộ nó bằng khung video và khung audio, và cung cấp video/audio/phụ đề đến màn hình video/audio/phụ đề 3216.

Sáng chế không bị giới hạn ở hệ thống được đề cập ở trên, và hoặc thiết bị mã hóa hình ảnh hoặc thiết bị giải mã hình ảnh trong các phương án được đề cập ở trên có thể được tích hợp trong hệ thống khác, ví dụ, hệ thống của xe hơi.

Các toán tử toán học

Các toán tử toán học được sử dụng trong sáng chế này tương tự với các toán tử được sử dụng trong ngôn ngữ lập trình C. Tuy nhiên, các kết quả của phép chia số nguyên và các phép toán độ dịch số học được xác định chính xác hơn, và các phép toán bổ sung được xác định, chẳng hạn như lũy thừa và phép chia giá trị

thực. Các quy ước đếm và tính thông thường bắt đầu từ 0, ví dụ, "thứ nhất" tương đương với thứ 0, "thứ hai" tương đương với thứ 1, vân vân.

Các toán tử số học

Các toán tử số học sau đây được xác định như sau:

$+$ Phép cộng.

$-$ Phép trừ (dưới dạng toán tử hai đối số) hoặc phủ định (dưới dạng toán tử tiền tố một ngôi).

$*$ Phép nhân, bao gồm phép nhân ma trận.

x^y Lũy thừa. Xác định x theo lũy thừa y . Trong các ngữ cảnh khác, ký hiệu như vậy được sử dụng để viết lên trên nhằm mục đích giải thích là lũy thừa.

$/$ Phép chia số nguyên cắt cụt kết quả về 0. Ví dụ, $7 / 4$ và $-7 / -4$ được cắt cụt về 1 và $-7 / 4$ và $7 / -4$ được cắt cụt về -1 .

$\div$ Được sử dụng để thể hiện phép chia trong các phương trình toán học trong đó không có sự cắt cụt hoặc làm tròn.

$\frac{x}{y}$ Được sử dụng để thể hiện phép chia trong các phương trình toán học trong đó không có sự cắt cụt hoặc làm tròn.

$\sum_{i=x}^y f(i)$ Phép lấy tổng của $f(i)$ với i nhận tất cả các giá trị nguyên từ x trở lên và bao gồm y .

$x \% y$ Phép toán modulo. Phần còn lại của x được chia cho y , được xác định chỉ đối với các số nguyên x và y với $x \geq 0$ và $y > 0$.

Các toán tử logic

Các toán tử logic sau đây được xác định như sau:

$x \&\& y$ Logic lý luận "và" của x và y

$x || y$ Logic lý luận "và" của x và y

! Logic lý luận "không"

$x ? y : z$ Nếu x là Đúng hoặc không bằng 0, thì đánh các giá trị của y; nếu không thì, đánh giá giá trị của z.

Các toán tử quan hệ

Các toán tử sau đây được xác định như sau:

> Lớn hơn

>= Lớn hơn hoặc bằng

< Nhỏ hơn

<= Nhỏ hơn hoặc bằng

= = Bằng

!= Không bằng

Khi toán tử quan hệ được áp dụng cho thành phần cú pháp hoặc biến mà được gán giá trị "na" (không áp dụng), giá trị "na" này được coi là giá trị khác biệt đối với thành phần cú pháp hoặc biến. Giá trị "na" được coi là sẽ không bằng với bất kỳ giá trị khác.

Các toán tử theo từng bit

Các toán tử theo từng bit sau đây được xác định như sau:

& Từng bit "và". Khi thực hiện các phép toán trên các đối số nguyên, thực hiện các phép toán trên phép biểu diễn phần bù giá trị nguyên của hai. Khi thực hiện các phép toán trên đối số nhị phân mà chứa ít bit hơn đối số khác, đối số ngắn hơn được mở rộng bằng cách bổ sung nhiều bit có nghĩa bằng 0.

| Từng bit "hoặc". Khi thực hiện các phép toán trên các đối số nguyên, thực hiện các phép toán trên phép biểu diễn phần bù giá trị nguyên của hai. Khi thực hiện các phép toán trên đối số nhị phân mà chứa ít

bit hơn đối số khác, đối số ngắn hơn được mở rộng bằng cách bổ sung nhiều bit có nghĩa bằng 0.

- ^ Từng bit "loại trừ hoặc". Khi thực hiện các phép toán trên các đối số nguyên, thực hiện các phép toán trên phép biểu diễn phần bù giá trị nguyên của hai. Khi thực hiện các phép toán trên đối số nhị phân mà chứa ít bit hơn đối số khác, đối số ngắn hơn được mở rộng bằng cách bổ sung nhiều bit có nghĩa bằng 0.
- $x \gg y$ Dịch chuyển sang phải số học của phép biểu diễn số nguyên bù của x bởi y các chữ số nhị phân của hai. Chức năng này được xác định chỉ cho các giá trị số nguyên không âm của y . Các bit được dịch chuyển trong các bit có nghĩa nhất (MSBs - most significant bits) là kết quả của dịch chuyển sang phải có giá trị bằng với MSB của x trước phép toán dịch chuyển.
- $x \ll y$ Dịch chuyển sang trái số học của phép biểu diễn số nguyên bù của x bởi y các chữ số nhị phân của hai. Chức năng này được xác định chỉ cho các giá trị số nguyên không âm của y . Các bit được dịch chuyển trong các bit có nghĩa tối thiểu (least significant bits - LSBs) là kết quả của dịch chuyển sang trái có giá trị bằng 0.

Các toán tử gán

Các toán tử số học sau đây được xác định như sau:

- = Toán tử gán
- ++ Lượng gia, tức là, $x++$ tương đương với $x = x + 1$; khi được sử dụng trong chỉ số mảng, đánh giá giá trị của biến trước phép toán lượng gia.
- Giảm lượng, tức là, $x--$ tương đương với $x = x - 1$; khi được sử dụng trong chỉ số mảng, đánh giá giá trị của biến trước phép toán giảm lượng.

$+=$ Lượng gia theo lượng được xác định, tức là, $x += 3$ tương đương với $x = x + 3$, và $x += (-3)$ tương đương với $x = x + (-3)$.

$-=$ Giảm lượng theo lượng được xác định, tức là, $x -= 3$ tương đương với $x = x - 3$, và $x -= (-3)$ tương đương với $x = x - (-3)$.

Ký hiệu khoảng

Ký hiệu sau đây được sử dụng để xác định các khoảng giá trị:

$x = y..z$ nhận các giá trị nguyên bắt đầu từ y đến z , bao gồm cả hai giá trị này, với x , y , và z là các số nguyên và z lớn hơn y .

Các hàm số học

Các hàm số học được xác định:

$$\text{Abs}(x) = \begin{cases} x & ; \quad x \geq 0 \\ -x & ; \quad x < 0 \end{cases}$$

$\text{Asin}(x)$ hàm sin ngược lượng giác, thực hiện các phép toán trên x mà nằm trong khoảng từ $-1,0$ đến $1,0$, bao gồm cả hai giá trị này, với giá trị đầu ra nằm trong khoảng từ $-\pi/2$ đến $\pi/2$, bao gồm cả hai giá trị này, theo các đơn vị radian

$\text{Atan}(x)$ hàm tang ngược lượng giác, thực hiện các phép toán trên đối số x , có giá trị đầu ra nằm trong khoảng từ $-\pi/2$ đến $\pi/2$, bao gồm cả hai giá trị này, theo các đơn vị radian

$$\text{Atan2}(y, x) = \begin{cases} \text{Atan}\left(\frac{y}{x}\right) & ; \quad x > 0 \\ \text{Atan}\left(\frac{y}{x}\right) + \pi & ; \quad x < 0 \ \&\& \ y \geq 0 \\ \text{Atan}\left(\frac{y}{x}\right) - \pi & ; \quad x < 0 \ \&\& \ y < 0 \\ +\frac{\pi}{2} & ; \quad x = 0 \ \&\& \ y \geq 0 \\ -\frac{\pi}{2} & ; \quad \text{trường hợp khác} \end{cases}$$

$\text{Ceil}(x)$ số nguyên nhỏ nhất lớn hơn hoặc bằng x .

$$\text{Clip1}_Y(x) = \text{Clip3}(0, (1 \ll \text{BitDepth}_Y) - 1, x)$$

$$\text{Clip1}_C(x) = \text{Clip3}(0, (1 \ll \text{BitDepth}_C) - 1, x)$$

$$\text{Clip3}(x, y, z) = \begin{cases} x & ; \quad z < x \\ y & ; \quad z > y \\ z & ; \quad \text{trường hợp khác} \end{cases}$$

$\text{Cos}(x)$ hàm cosin lượng giác thực hiện các phép toán trên đối số x theo các đơn vị radian.

$\text{Floor}(x)$ số nguyên lớn nhất nhỏ hơn hoặc bằng x .

$$\text{GetCurrMsb}(a, b, c, d) = \begin{cases} c + d & ; \quad b - a \geq d / 2 \\ c - d & ; \quad a - b > d / 2 \\ c & ; \quad \text{trường hợp khác} \end{cases}$$

$\text{Ln}(x)$ lôgarit tự nhiên của x (lôgarit cơ số e , trong đó e là cơ số lôgarit tự nhiên không đôi 2,718 281 828...).

$\text{Log2}(x)$ lôgarit cơ số 2 x .

$\text{Log10}(x)$ lôgarit cơ số 10 của x .

$$\text{Min}(x, y) = \begin{cases} x & ; \quad x \leq y \\ y & ; \quad x > y \end{cases}$$

$$\text{Max}(x, y) = \begin{cases} x & ; \quad x \geq y \\ y & ; \quad x < y \end{cases}$$

$$\text{Round}(x) = \text{Sign}(x) * \text{Floor}(\text{Abs}(x) + 0,5)$$

$$\text{Sign}(x) = \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x == 0 \\ -1 & ; \quad x < 0 \end{cases}$$

$\text{Sin}(x)$ Hàm sin lượng giác thực hiện các phép toán trên đối số x theo các đơn vị radian

$$\text{Sqrt}(x) = \sqrt{x}$$

$$\text{Swap}(x, y) = (y, x)$$

$\text{Tan}(x)$ hàm tang lượng giác thực hiện các phép toán trên đối số x theo các đơn vị radian

Thứ tự ưu tiên của các phép toán

Khi thứ tự ưu tiên không được thể hiện rõ ràng bằng sử dụng các dấu ngoặc đơn, áp dụng các quy tắc sau đây:

- Các phép toán có mức ưu tiên cao hơn được đánh giá trước bất kỳ phép toán nào có mức ưu tiên thấp hơn.
- Các phép toán có cùng mức ưu tiên được đánh giá tuần tự từ trái qua phải.

Bảng dưới đây xác định mức ưu tiên của các phép toán từ cao nhất đến thấp nhất; vị trí cao hơn trong bảng thể hiện mức ưu tiên cao hơn.

Đối với các toán tử đó mà cũng được sử dụng trong ngôn ngữ lập trình C, thứ tự ưu tiên được sử dụng trong bản mô tả này giống như được sử dụng trong ngôn ngữ lập trình C.

Bảng: mức ưu tiên của phép toán từ cao nhất (ở đỉnh bảng) đến thấp nhất (ở cuối bản)

Các phép toán (với các toán hạng x, y, và z)
"x++", "x--"
"!x", "-x" (là toán tử tiền tố một ngôi)
x^y
"x * y", "x / y", "x ÷ y", " $\frac{x}{y}$ ", "x % y"
"x + y", "x - y" (là toán tử hai đối số), " $\sum_{i=x}^y f(i)$ "
"x << y", "x >> y"
"x < y", "x <= y", "x > y", "x >= y"
"x == y", "x != y"
"x & y"
"x y"
"x && y"
"x y"
"x ? y : z"
"x..y"
"x = y", "x += y", "x -= y"

Văn bản mô tả các toán tử logic

Trong văn bản này, câu lệnh của các toán tử logic là như sẽ được mô tả về mặt toán học theo dạng sau:

if(điều kiện 0)

 câu lệnh 0

else if(điều kiện 1)

 câu lệnh 1

...

else /* chú thích lại thông tin về các điều kiện còn lại */
câu lệnh n

có thể được mô tả theo cách sau:

... như sau / ... sau đây áp dụng:

- If điều kiện 0, câu lệnh 0
- Nếu không thì, if điều kiện 1, câu lệnh 1
- ...
- Nếu không thì (chú thích thông tin về điều kiện còn lại), thì thực hiện câu lệnh n

Mỗi "If ... Nếu không thì, if ... Nếu không thì, ..." câu lệnh trong văn bản này được đưa vào "... như sau " hoặc "... sau đây áp dụng" ngay sau đó là "If ... ". Điều kiện cuối cùng của "If ... Nếu không thì, if ... Nếu không thì, ..." luôn là "Nếu không thì, ...". Xen vào giữa "If ... Nếu không thì, if ... Nếu không thì, ..." các câu lệnh có thể được xác nhận bằng cách so khớp "... như sau" hoặc "... sau đây áp dụng" với kết thúc là "Nếu không thì, ...".

Trong văn bản này, câu lệnh của các phép toán logic là như sẽ được mô tả về mặt toán học theo dạng sau đây:

```
if( điều kiện 0a && điều kiện 0b )
    câu lệnh 0
else if( điều kiện 1a || điều kiện 1b )
    câu lệnh 1
...
else
    câu lệnh n
```

có thể được mô tả theo cách sau:

... như sau / ... sau đây áp dụng:

- If tất cả các điều kiện sau đây là đúng, câu lệnh 0:

- điều kiện 0a
- điều kiện 0b
- Nếu không thì, nếu một hoặc nhiều các điều kiện sau đây là đúng, câu lệnh 1:
 - điều kiện 1a
 - điều kiện 1b
- ...
- Nếu không thì, câu lệnh n

Trong văn bản này, câu lệnh của các phép toán logic là như sẽ được mô tả về mặt toán học theo dạng sau đây:

if(điều kiện 0)

 câu lệnh 0

if(điều kiện 1)

 câu lệnh 1

có thể được mô tả theo cách sau:

Khi điều kiện 0, câu lệnh 0

Khi điều kiện 1, câu lệnh 1

Mặc dù các phương án của sáng chế được mô tả sơ bộ dựa trên mã hóa video, nhưng cần lưu ý rằng các phương án về hệ thống lập mã 10, bộ mã hóa 20 và bộ giải mã 30 (và hệ thống 10 một cách tương ứng) và các phương án được mô tả ở đây có thể còn được tạo cấu hình vẫn để xử lý hình ảnh hoặc mã hóa, tức là việc xử lý hoặc mã hóa hình ảnh riêng độc lập với bất kỳ hình ảnh trước đó hoặc tiếp theo khi mã hóa video. Nói chung chỉ các bộ dự báo nội ảnh 244 (bộ mã hóa) và 344 (bộ giải mã) có thể không khả dụng trong trường hợp mã hóa xử lý hình ảnh bị giới hạn ở một hình ảnh 17. Tất cả các chức năng khác (còn được gọi là các công cụ hoặc các kỹ thuật) của bộ mã hóa video 20 và bộ giải mã video 30 có thể được sử dụng như nhau vẫn để xử lý hình ảnh, ví dụ tính toán phần dư 204/304,

biến đổi 206, lượng tử hóa 208, lượng tử hóa ngược 210/310, biến đổi (ngược) 212/312, phân chia 262/362, dự báo nội ảnh 254/354, và/hoặc lọc vòng 220, 320, và mã hóa entropi 270 và giải mã entropi 304.

Các phương án, ví dụ về bộ mã hóa 20 và bộ giải mã 30, và các chức năng được mô tả ở đây, ví dụ dựa vào bộ mã hóa 20 và bộ giải mã 30, có thể được thực hiện trong phần cứng, phần mềm, vi chương trình, hoặc bất kỳ kết hợp nào của chúng. Nếu được thực hiện trong phần mềm, thì các chức năng này có thể được lưu trữ trên phương tiện có thể đọc được bằng máy tính hoặc được truyền qua đa phương tiện truyền thông dưới dạng một hoặc nhiều lệnh hoặc mã và được thực hiện bởi bộ xử lý dựa trên phần cứng. Phương tiện có thể đọc được bằng máy tính có thể bao gồm phương tiện lưu trữ có thể đọc được bằng máy tính, mà tương ứng với phương tiện hữu hình chẳng hạn như phương tiện lưu trữ dữ liệu, hoặc phương tiện truyền thông bao gồm bất kỳ phương tiện nào thuận lợi để chuyển chương trình máy tính từ một vị trí đến vị trí khác, ví dụ, theo giao thức truyền thông. Trong cách thức này, phương tiện có thể đọc được bằng máy tính thông thường có thể tương ứng với (1) phương tiện lưu trữ có thể đọc được bằng máy tính hữu hình mà là phương tiện bất biến hoặc phương tiện truyền thông (2) chẳng hạn như tín hiệu hoặc sóng mang. Phương tiện lưu trữ dữ liệu có thể là phương tiện khả dụng bất kỳ mà có thể được truy nhập bởi một hoặc nhiều máy tính hoặc một hoặc nhiều bộ xử lý để truy lục các lệnh, mã và/hoặc các cấu trúc dữ liệu để thực hiện các kỹ thuật được mô tả trong sáng chế. Sản phẩm chương trình máy tính có thể bao gồm phương tiện có thể đọc được bằng máy tính.

Cụ thể, sáng chế đề xuất phương pháp giải mã dòng bit video được lập mã được thực hiện trong bộ giải mã như được minh họa trên FIG.9, phương pháp này bao gồm: S901, thu thành phần cú pháp mức tham số chuỗi (SPS) từ dòng bit, trong đó thành phần cú pháp mức SPS này bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS. S902, thu, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp xác định liệu một hoặc nhiều

hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không; và S903, dự báo một hoặc nhiều hình ảnh được lập mã dựa trên giá trị của thành phần cú pháp cho phép liên lớp.

Tương tự, sáng chế đề xuất phương pháp mã hóa dòng bit video bao gồm dữ liệu được lập mã để được thực hiện trong bộ mã hóa như được minh họa trên FIG.10, phương pháp này bao gồm: S1001, mã hóa thành phần cú pháp mức tập tham số chuỗi (SPS) trong dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; S1003, mã hóa, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp trong dòng bit, trong đó thành phần cú pháp cho phép liên lớp này xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không.

Ngoài ra, phương pháp có thể còn bao gồm: S1002, xác định liệu thành phần cú pháp mức SPS có lớn hơn giá trị thiết đặt trước hay không.

FIG.11 minh họa bộ giải mã 1100 được tạo cấu hình để giải mã dòng bit video bao gồm dữ liệu được lập mã dùng cho nhiều hình ảnh. Bộ giải mã 1100 theo ví dụ được thể hiện bao gồm: bộ thu 1110 được tạo cấu hình để thu thành phần cú pháp mức tập tham số chuỗi (SPS) từ dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; bộ thu 1110 còn được tạo cấu hình để thu, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không; và bộ

dự báo 1120 được tạo cấu hình để dự báo một hoặc nhiều hình ảnh được lập mã dựa trên giá trị của thành phần cú pháp cho phép liên lớp.

Trong đó bộ thu có thể là bộ giải mã entrôpi 304. Bộ dự báo 1120 có thể là bộ dự báo liên ảnh 344. Bộ giải mã 1100 có thể là thiết bị đích 14, bộ giải mã 30, thiết bị 500, bộ giải mã video 3206 hoặc thiết bị đầu cuối 3106.

Tương tự, sáng chế đề xuất bộ mã hóa 1200 được tạo cấu hình dùng để mã hóa dòng bit video bao gồm dữ liệu được lập mã cho nhiều hình ảnh như được minh họa trên FIG.12. Bộ mã hóa 1200 bao gồm: bộ mã hóa thứ nhất 1210 được tạo cấu hình để mã hóa thành phần cú pháp mức tập tham số chuỗi (SPS) trong dòng bit, trong đó thành phần cú pháp mức SPS bằng với giá trị thiết đặt trước xác định rằng không có tập tham số video (VPS) nào được tham chiếu đến bởi SPS, và thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước xác định rằng SPS tham chiếu đến VPS; bộ mã hóa thứ hai 1220 được tạo cấu hình để mã hóa, dưới dạng thành phần cú pháp mức SPS lớn hơn giá trị thiết đặt trước, thành phần cú pháp cho phép liên lớp trong dòng bit, trong đó thành phần cú pháp cho phép liên lớp này xác định liệu một hoặc nhiều hình ảnh tham chiếu liên lớp (ILRP) có được cho phép để sẽ được sử dụng cho dự báo liên ảnh của một hoặc nhiều hình ảnh được lập mã hay không.

Trong các cách thức thực hiện khả thi của phương pháp theo khía cạnh thứ tư như vậy, bộ mã hóa còn bao gồm bộ xác định được tạo cấu hình để xác định liệu thành phần cú pháp mức SPS có lớn hơn giá trị thiết đặt trước hay không.

Trong đó bộ mã hóa thứ nhất 1210 và bộ mã hóa thứ hai 1220 có thể là bộ mã hóa entrôpi 270. Bộ xác định có thể là bộ lựa chọn chế độ 260. Bộ mã hóa 1200 có thể là thiết bị nguồn 12, bộ mã hóa 20, hoặc thiết bị 500.

Ví dụ, và không bị giới hạn, phương tiện lưu trữ có thể đọc được bằng máy tính như vậy có thể bao gồm RAM, ROM, EEPROM, CD-ROM hoặc các thiết bị lưu trữ đĩa quang khác, thiết bị lưu trữ đĩa từ, hoặc các thiết bị lưu trữ từ khác, bộ nhớ chớp, hoặc phương tiện bất kỳ khác mà có thể được sử dụng để lưu trữ mã chương trình mong muốn dưới dạng các lệnh hoặc các cấu trúc dữ liệu và có thể được truy

nhập bằng máy tính. Ngoài ra, bất kỳ kết nối nào được gọi đúng là phương tiện có thể đọc được bằng máy tính. Ví dụ, nếu các lệnh được truyền từ trang mạng, máy chủ, hoặc các nguồn từ xa khác bằng cách sử dụng cáp đồng trục, cáp sợi quang, cáp xoắn hai sợi, đường thuê bao số (digital subscriber line - DSL), hoặc các kỹ thuật không dây khác chẳng hạn như hồng ngoại, vô tuyến, và vi sóng, sau đó cáp đồng trục, cáp sợi quang, cáp xoắn hai sợi, DSL, hoặc các kỹ thuật truyền thông khác chẳng hạn như hồng ngoại, vô tuyến, và vi sóng được bao gồm trong định nghĩa về phương tiện. Cần hiểu rằng, tuy nhiên, phương tiện lưu trữ có thể đọc được bằng máy tính và phương tiện lưu trữ dữ liệu không bao gồm các kết nối, các sóng mang, báo hiệu, hoặc các phương tiện khả biến khác, mà thay vào đó hướng đến phương tiện lưu trữ bất biến, hữu hình. Ổ đĩa và đĩa, như được sử dụng ở đây, bao gồm đĩa compact (CD), đĩa laze, đĩa quang học, đĩa đa dụng số (DVD), đĩa mềm và đĩa Blu-ray, trong đó các đĩa thường được sao chép dữ liệu một cách từ tính, trong khi các đĩa sao chép dữ liệu một cách quang học bằng laze. Các kết hợp ở trên có thể sẽ được bao gồm trong phạm vi của phương tiện có thể đọc được bằng máy tính.

Các lệnh có thể được thực hiện bởi một hoặc nhiều bộ xử lý, chẳng hạn như một hoặc nhiều bộ xử lý tín hiệu kỹ thuật số (DSP - digital signal processors), các bộ vi xử lý mục đích chung, các mạch tích hợp ứng dụng cụ thể (ASICs), các mảng cổng lập trình được dạng trường (FPGAs), hoặc hệ mạch logic rời rạc được tích hợp tương đương khác. Do đó, thuật ngữ “bộ xử lý”, như được sử dụng ở đây có thể đề cập đến cấu trúc trước đó hoặc bất kỳ cấu trúc nào phù hợp để thực hiện các kỹ thuật được mô tả ở đây. Ngoài ra, trong một vài khía cạnh, chức năng được mô tả ở đây có thể bố trí trong các môđun phần cứng và/hoặc phần mềm chuyên dụng được tạo cấu hình dùng để mã hóa và giải mã, hoặc được tích hợp trong bộ mã hóa và giải mã kết hợp. Ngoài ra, các kỹ thuật có thể hoàn toàn được thực hiện trong một hoặc nhiều mạch có các thành phần logic.

Các kỹ thuật của sáng chế này có thể được thực hiện trong một loạt các thiết bị hoặc các thiết bị, bao gồm thiết bị cầm tay không dây, mạch tích hợp (IC - integrated circuit) hoặc tập các IC (ví dụ, bộ vi mạch). Các thành phần, các môđun,

hoặc các bộ khác nhau được mô tả trong sáng chế này để nhấn mạnh các khía cạnh chức năng của các thiết bị được tạo cấu hình để thực hiện các kỹ thuật được bộc lộ, mà không nhất thiết yêu cầu được nhận biết bởi các đơn vị phần cứng khác nhau. Thay vào đó, như được mô tả ở trên, các bộ khác nhau có thể được kết hợp trong bộ phận phần cứng mã hóa và giải mã hoặc được bố trí bởi tập hợp các bộ phận phần cứng liên hoạt động, bao gồm một hoặc nhiều bộ xử lý như được mô tả ở trên, kết hợp với phần mềm và/hoặc vi chương trình phù hợp.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã dòng bit video được lập mã, phương pháp này bao gồm:

thu nhận phần tử cú pháp mức tập tham số chuỗi (SPS - sequence parameter set) từ dòng bit video được lập mã, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước chỉ báo rằng không có tập tham số video (VPS - video parameter set) được viện dẫn bởi SPS, và trong đó phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước chỉ báo rằng SPS viện dẫn tới VPS;

thu nhận, từ dòng bit video được lập mã, phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ảnh tham chiếu liên lớp (ILRP - inter-layer reference picture) có được cho phép sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã hay không khi phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước; và

thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã dựa trên giá trị của phần tử cú pháp được cho phép liên lớp.

2. Phương pháp theo điểm 1, trong đó VPS bao gồm nhiều phần tử cú pháp mà mô tả thông tin dự đoán liên lớp của các lớp trong chuỗi video được mã hóa (CVS - coded video sequence) bao gồm một hoặc nhiều ILRP và một hoặc nhiều ảnh được lập mã, và trong đó SPS bao gồm phần tử cú pháp mức SPS và phần tử cú pháp được cho phép liên lớp.

3. Phương pháp theo điểm 2, trong đó bước thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã dựa trên giá trị của phần tử cú pháp được cho phép liên lớp bao gồm:

thu nhận một hoặc nhiều ILRP dựa trên thông tin dự đoán liên lớp; và

thực hiện thêm dự đoán liên đới trên một hoặc nhiều ảnh lập được mã bằng cách viện dẫn một hoặc nhiều ILRP khi phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ILRP được cho phép sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã.

4. Phương pháp theo điểm 1, trong đó ảnh được lập mã trong số một hoặc nhiều ảnh được lập mã và ILRP của ảnh được lập mã thuộc về các lớp khác nhau.

5. Phương pháp theo điểm 1, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước còn chỉ báo rằng chuỗi video được mã hóa (CVS - coded video sequence) bao gồm các ảnh được lập mã của chỉ một lớp.

6. Phương pháp theo điểm 1, trong đó giá trị được thiết lập trước bằng 0.

7. Phương pháp theo điểm 1, trong đó bước thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã dựa trên giá trị của phần tử cú pháp được cho phép liên lớp bao gồm dự đoán một hoặc nhiều ảnh được lập mã mà không vi phạm bất kỳ ILRP khi phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ILRP không được sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã.

8. Phương pháp mã hóa dòng bit video được lập mã, phương pháp này bao gồm:

mã hóa phần tử cú pháp mức tập tham số chuỗi (SPS - sequence parameter set) thành dòng bit video được lập mã, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước chỉ báo rằng không có tập tham số video (VPS - video parameter set) được vi phạm bởi SPS, và trong đó phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước chỉ báo rằng SPS vi phạm dẫn tới VPS;

mã hóa phần tử cú pháp được cho phép liên lớp thành dòng bit video được lập mã khi phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước,

trong đó phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ảnh tham chiếu liên lớp (ILRP - inter-layer reference picture) được cho phép sử dụng để thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã hay không.

9. Phương pháp theo điểm 8, trong đó VPS bao gồm nhiều phần tử cú pháp mà mô tả thông tin dự đoán liên lớp của các lớp trong chuỗi video được mã hóa (CVS - coded video sequence) bao gồm một hoặc nhiều ILRP và một hoặc nhiều ảnh được lập mã, và trong đó SPS bao gồm phần tử cú pháp mức SPS và phần tử cú pháp được cho phép liên lớp.

10. Phương pháp theo điểm 8, trong đó ảnh được lập mã trong số một hoặc nhiều ảnh được lập mã và ILRP của ảnh được lập mã thuộc về các lớp khác nhau.

11. Phương pháp theo điểm 8, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước còn chỉ báo rằng chuỗi video được mã hóa (CVS - coded video sequence) bao gồm các ảnh được lập mã của chỉ một lớp.

12. Phương pháp theo điểm 8, trong đó giá trị được thiết lập trước bằng 0.

13. Phương pháp theo điểm 8, trong đó bước mã hóa phần tử cú pháp được cho phép liên lớp thành dòng bit bao gồm mã hóa phần tử cú pháp được cho phép liên lớp mà chỉ báo rằng một hoặc nhiều ILRP được cho phép sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã thành dòng bit video được lập mã dựa trên một hoặc nhiều ILRP được cho phép sử dụng để thực hiện dự đoán

liên đới của một hoặc nhiều ảnh được lập mã.

14. Phương pháp theo điểm 8, trong đó bước mã hóa phần tử cú pháp được cho phép liên lớp thành dòng bit bao gồm mã hóa phần tử cú pháp được cho phép liên lớp mà chỉ báo rằng một hoặc nhiều ILRP không được sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã thành dòng bit video được lập mã dựa trên một hoặc nhiều ILRP không được cho phép sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã.

15. Thiết bị giải mã, bao gồm:

bộ nhớ được tạo cấu hình để lưu trữ các lệnh; và

bộ xử lý được ghép nối tới bộ nhớ và được tạo cấu hình để thực thi các lệnh, mà làm cho bộ xử lý được tạo cấu hình để:

thu nhận phần tử cú pháp mức tập tham số chuỗi (SPS - sequence parameter set) từ dòng bit video được lập mã, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước chỉ báo rằng không có tập tham số video (VPS - video parameter set) được viện dẫn bởi SPS, và trong đó phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước chỉ báo rằng SPS viện dẫn tới VPS;

thu nhận, từ dòng bit video được lập mã, phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ảnh tham chiếu liên lớp (ILRP - inter-layer reference picture) có được cho phép sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã hay không khi phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước; và

thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã dựa trên giá trị của phần tử cú pháp được cho phép liên lớp.

16. Thiết bị mã hóa, bao gồm

bộ nhớ được tạo cấu hình để lưu trữ các lệnh; và

bộ xử lý được ghép nối tới bộ nhớ và được tạo cấu hình để thực thi các lệnh, mà làm cho bộ xử lý được tạo cấu hình để:

mã hóa phần tử cú pháp mức tập tham số chuỗi (SPS - sequence parameter set) thành dòng bit video được lập mã, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước chỉ báo rằng không có tập tham số video (VPS - video parameter set) được viện dẫn bởi SPS, và trong đó phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước chỉ báo rằng SPS viện dẫn tới VPS; và

mã hóa phần tử cú pháp được cho phép liên lớp thành dòng bit video được lập mã khi phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước, trong đó phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ảnh tham chiếu liên lớp (ILRP) có được cho phép để thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã hay không.

17. Phương tiện bất biến đọc được bởi máy tính lưu trữ mã chương trình mà, khi được thực thi bởi máy tính hoặc bộ xử lý, làm cho máy tính hoặc bộ xử lý:

thu nhận phần tử cú pháp mức tập tham số chuỗi (SPS - sequence parameter set) từ dòng bit video được lập mã, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước chỉ báo rằng không có tập tham số video (VPS - video parameter set) được viện dẫn bởi SPS, và trong đó phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước chỉ báo rằng SPS viện dẫn tới VPS;

thu nhận, từ dòng bit video được lập mã, phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ảnh tham chiếu liên lớp (ILRP - inter-layer reference picture) có được cho phép sử dụng để thực hiện dự đoán liên đới của một hoặc nhiều ảnh được lập mã hay không khi phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước; và

thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã dựa trên giá trị của phần tử cú pháp được cho phép liên lớp.

18. Phương tiện bất biến đọc được bởi máy tính lưu trữ mã chương trình mà, khi được thực thi bởi máy tính hoặc bộ xử lý, làm cho máy tính hoặc bộ xử lý:

mã hóa phần tử cú pháp mức tập tham số chuỗi (SPS - sequence parameter set) thành dòng bit video được lập mã, trong đó phần tử cú pháp mức SPS bằng giá trị được thiết lập trước chỉ báo rằng không có tập tham số video (VPS - video parameter set) được viện dẫn bởi SPS, và trong đó phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước chỉ báo rằng SPS viện dẫn tới VPS; và

mã hóa phần tử cú pháp được cho phép liên lớp thành dòng bit video được lập mã khi phần tử cú pháp mức SPS lớn hơn giá trị được thiết lập trước, trong đó phần tử cú pháp được cho phép liên lớp chỉ báo rằng một hoặc nhiều ảnh tham chiếu liên lớp (ILRP) có được cho phép để thực hiện dự đoán liên đới trên một hoặc nhiều ảnh được lập mã hay không.

1/13

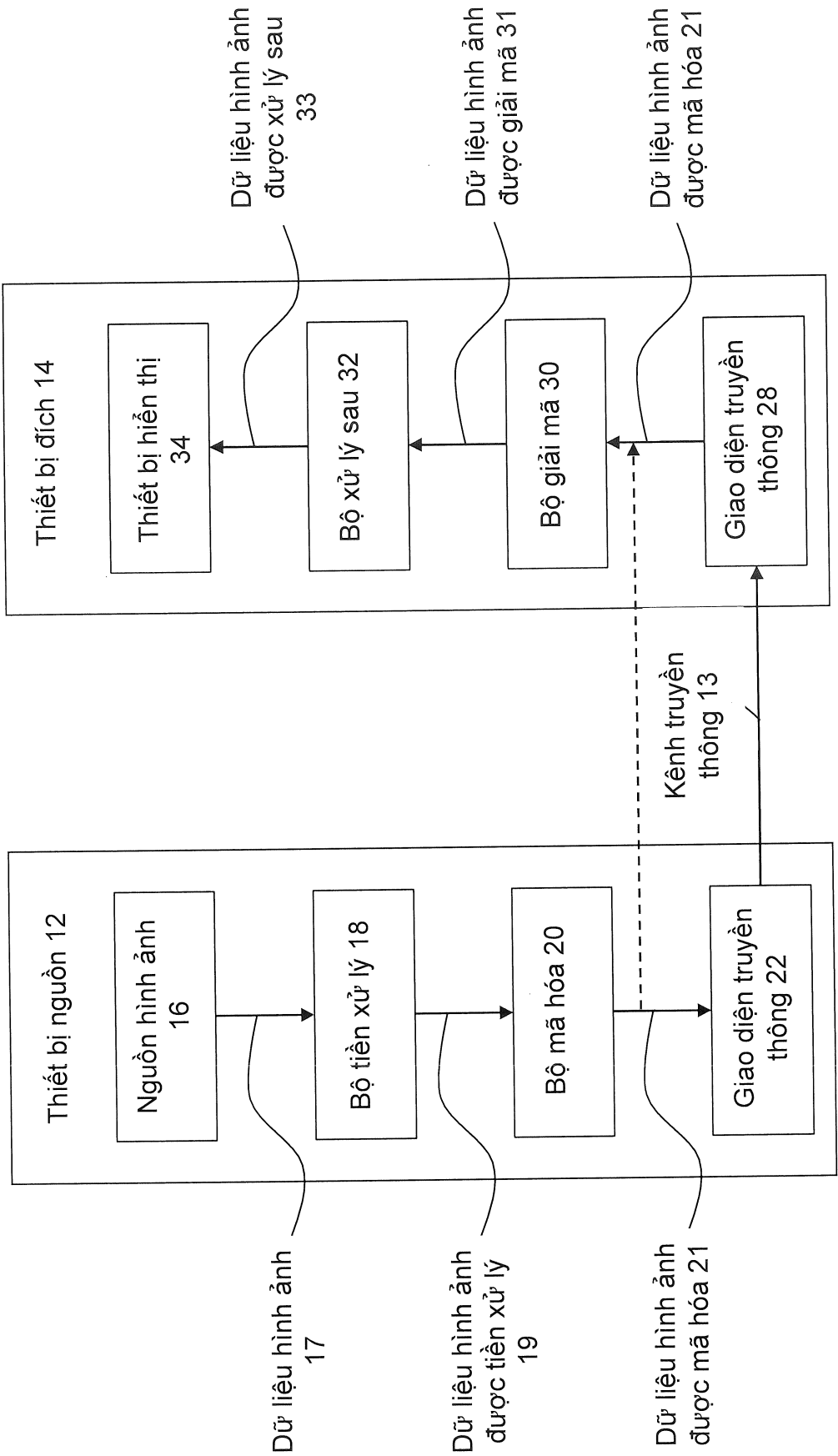


FIG. 1A

2/13

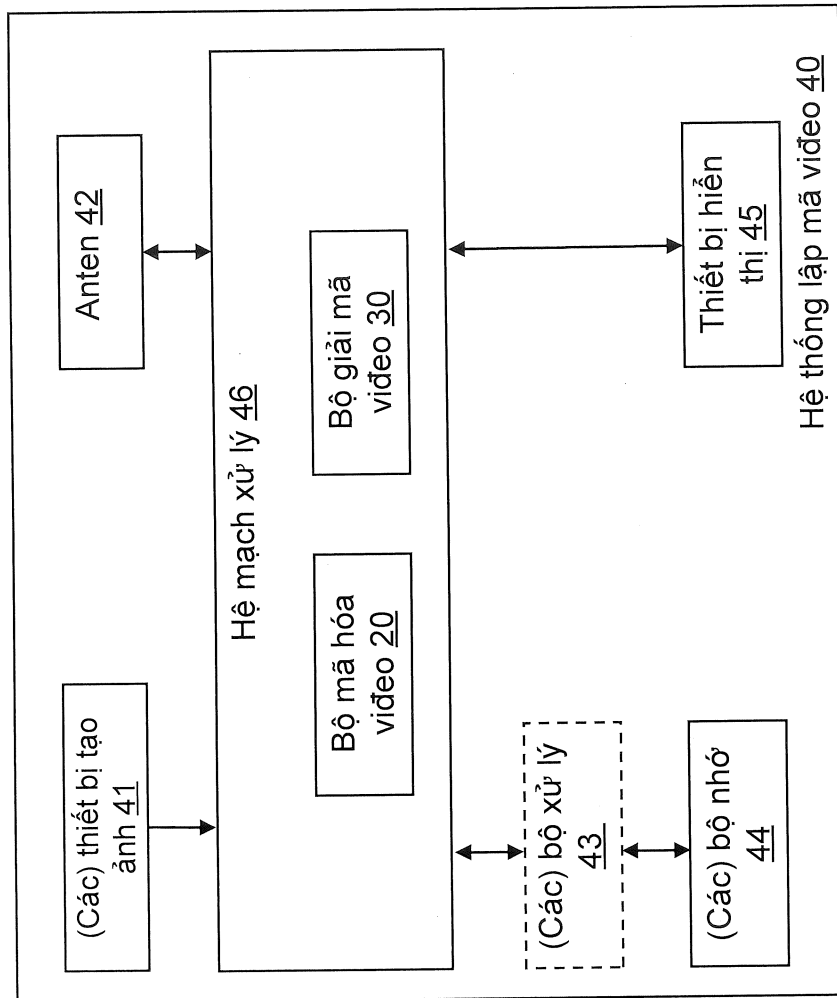


FIG. 1B

3/13

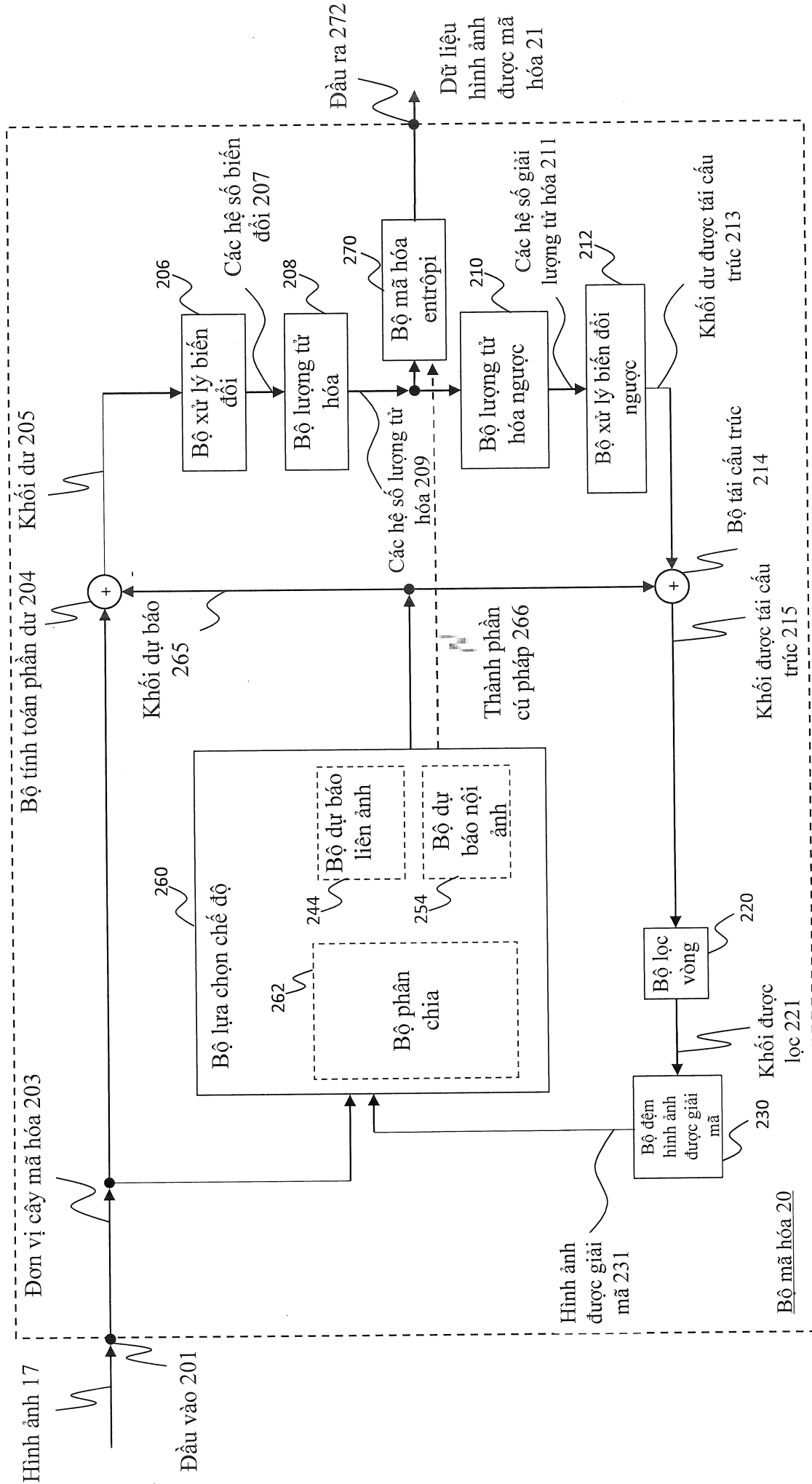


FIG. 2

4/13

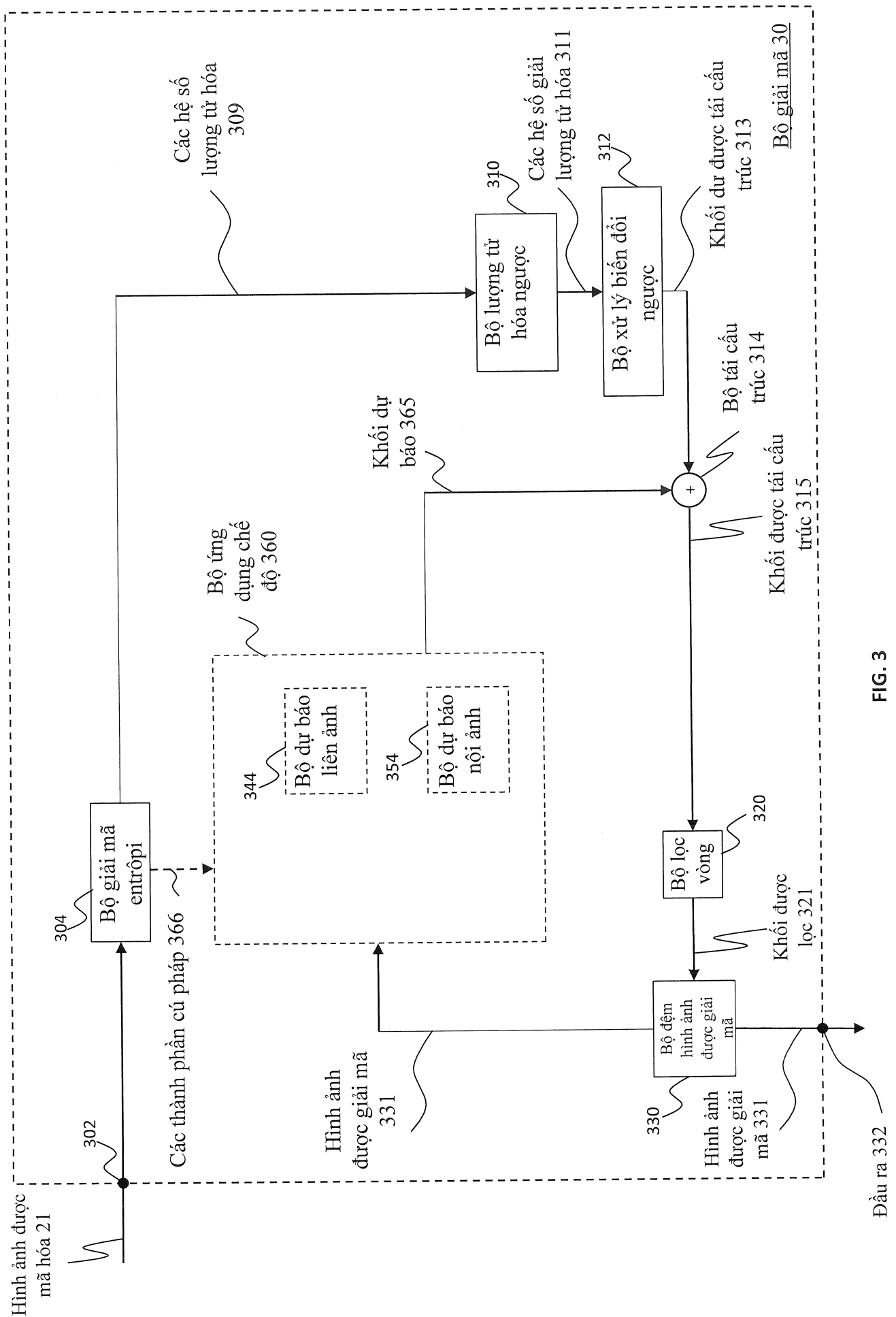
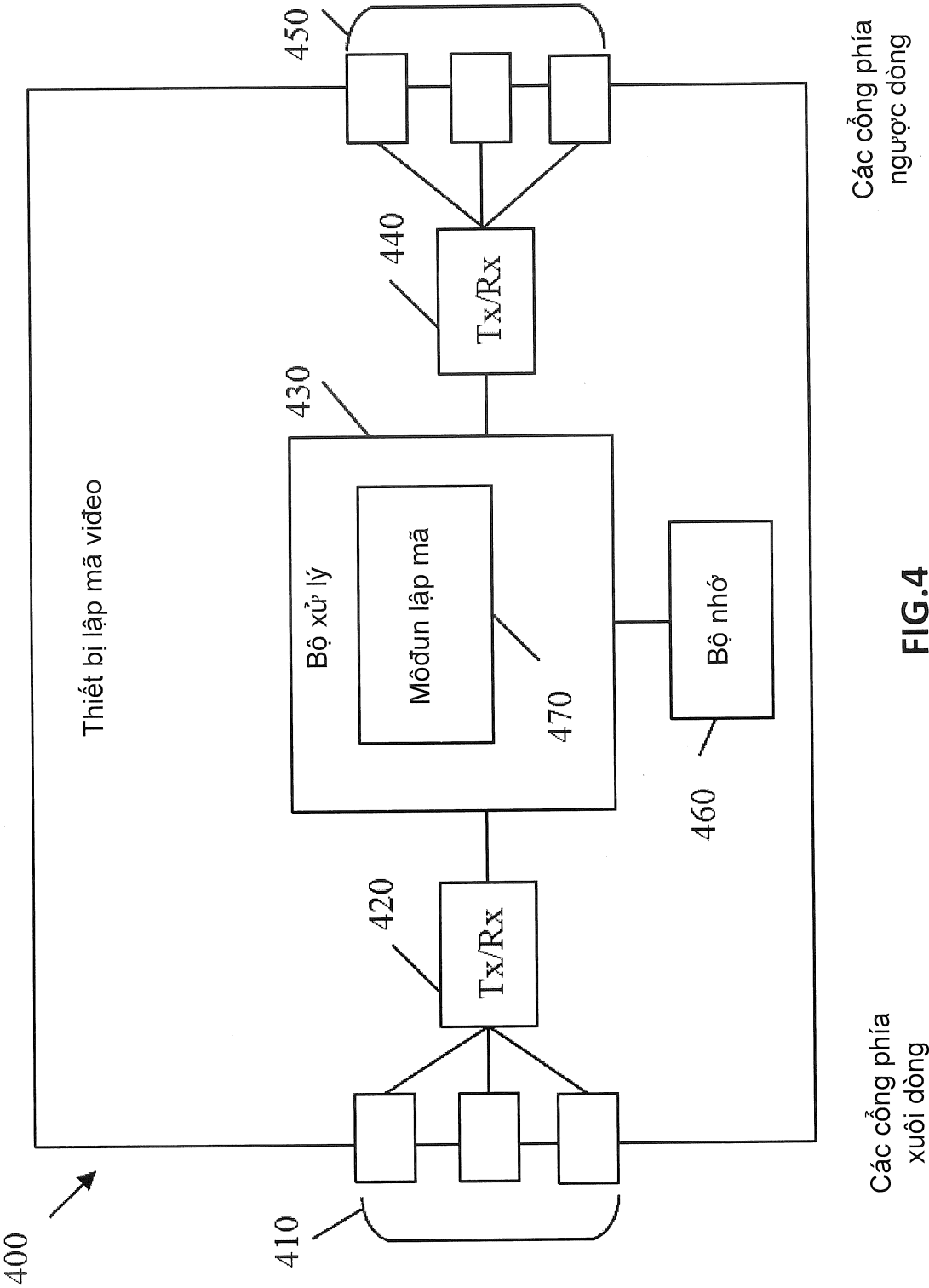


FIG. 3



6/13

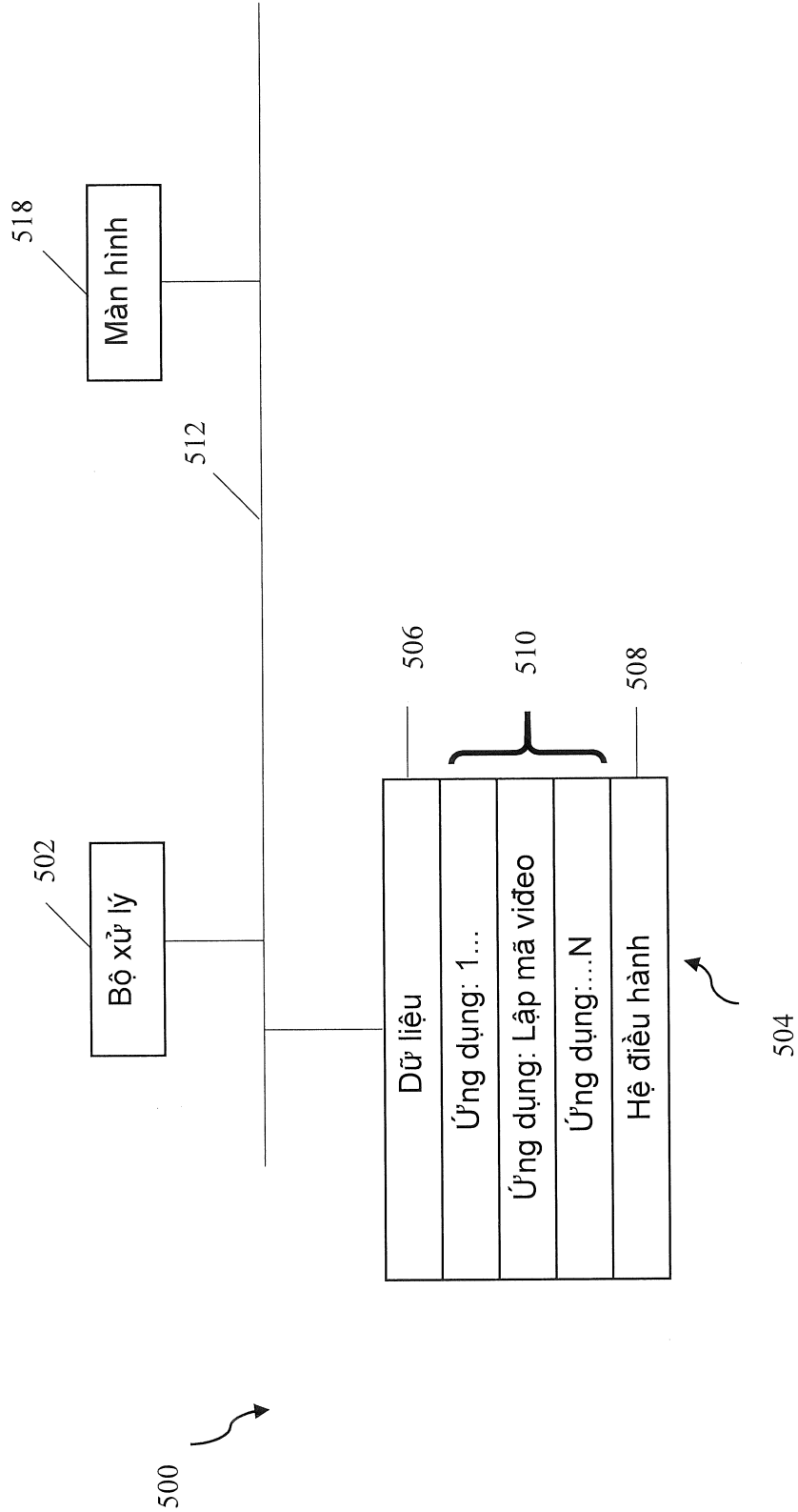


FIG. 5

7/13

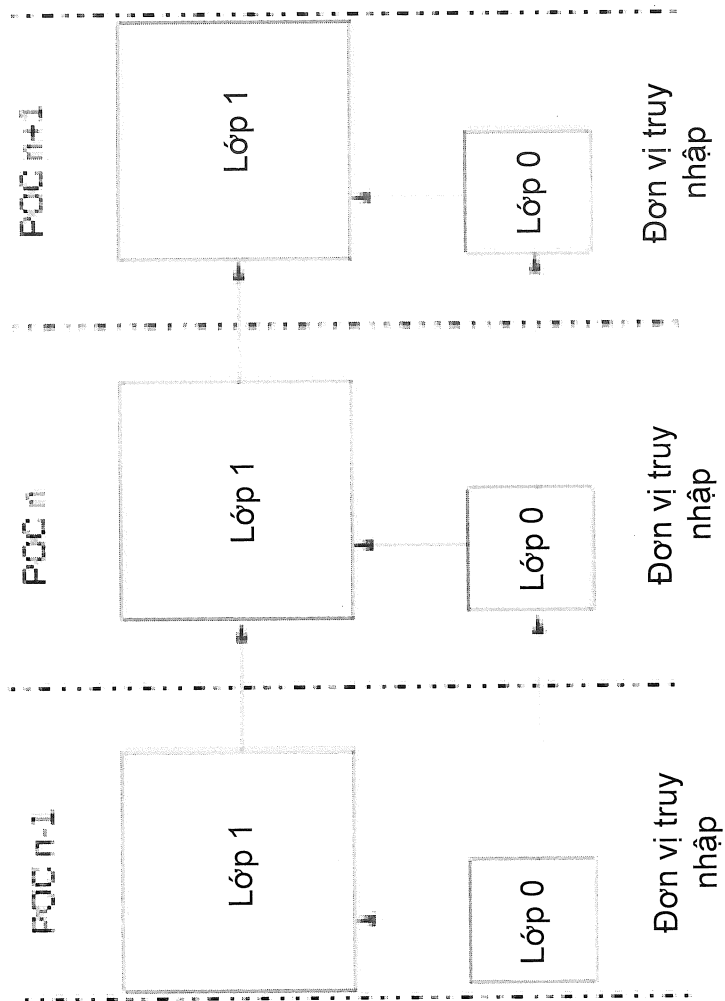


FIG. 6

8/13

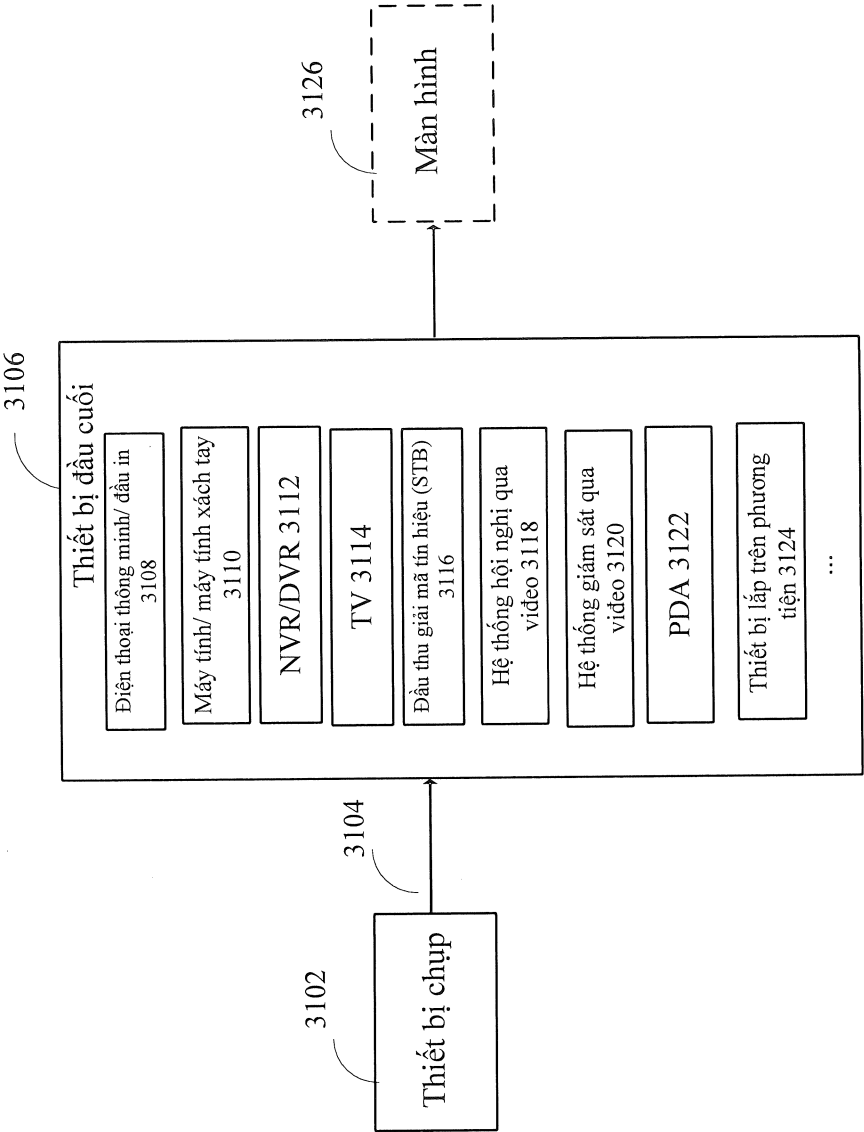


FIG. 7

9/13

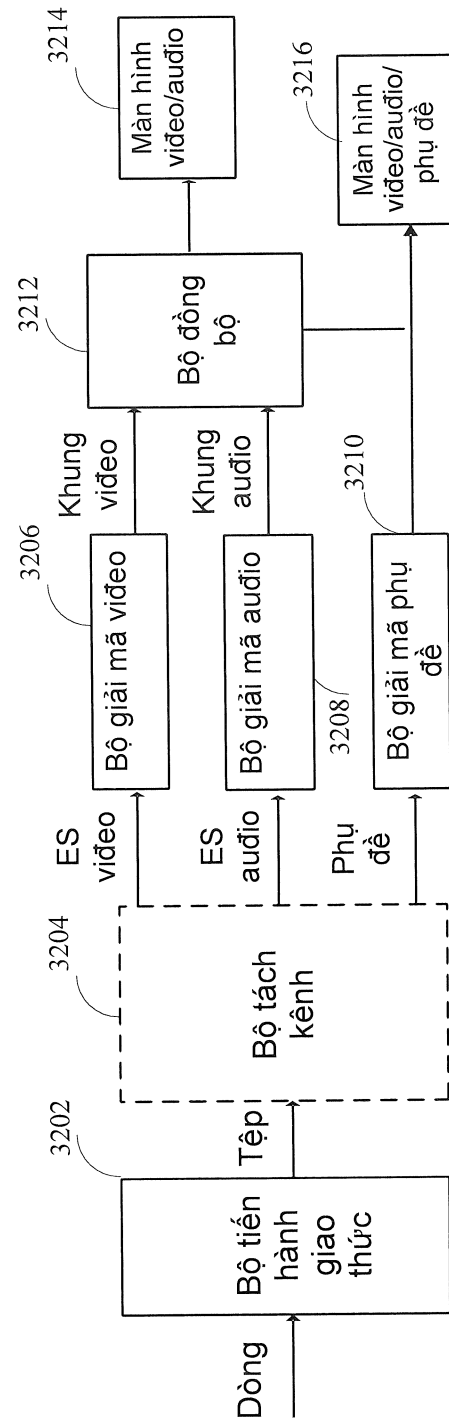


FIG. 8

10/13

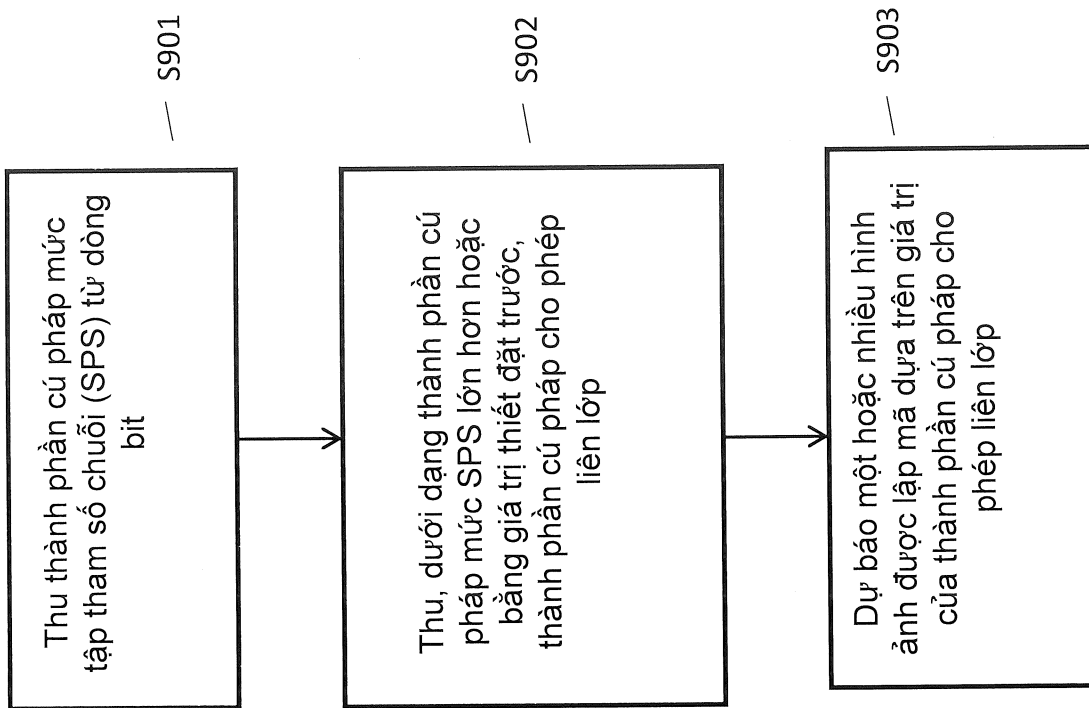


FIG. 9

11/13

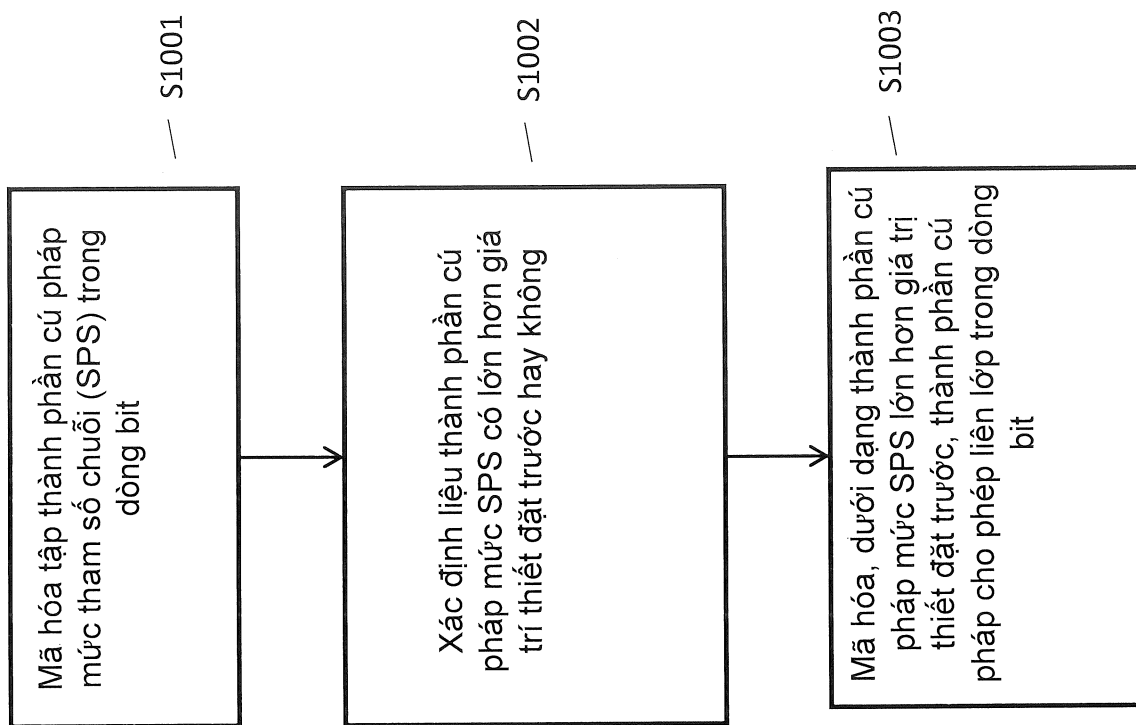


FIG. 10

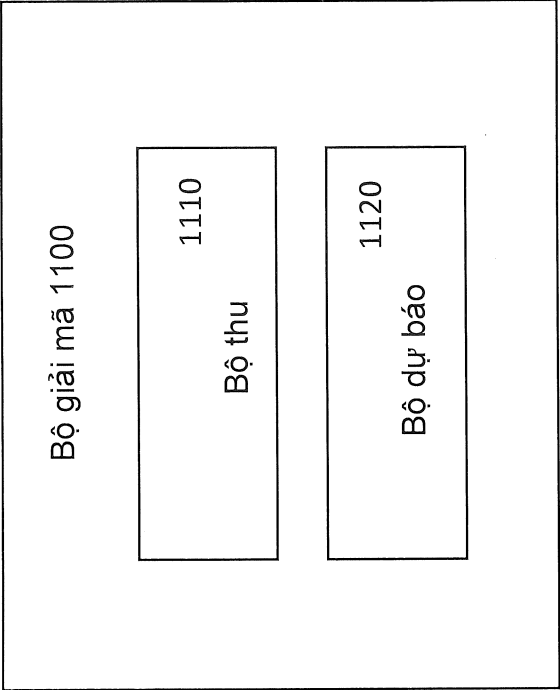


FIG. 11

13/13

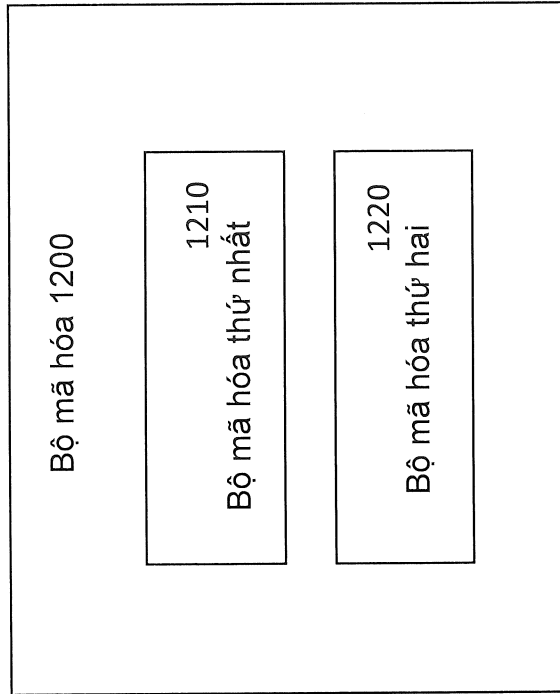


FIG. 12