



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẢNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0052148

(51)^{2020.01} H04N 19/50

(13) B

(21) 1-2022-02549

(22) 16/09/2020

(86) PCT/US2020/050985 16/09/2020

(87) WO2021/061462 01/04/2021

(30) 62/905,141 24/09/2019 US

(45) 27/10/2025 451

(43) 25/07/2022 412A

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

Huawei Administration Building, Bantian, Longgang District, Shenzhen, Guangdong
518129, China

(72) WANG, Ye-Kui (US).

(74) Công ty Luật TNHH T&G (TGVN)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ ĐỀ MÃ HÓA VÀ GIẢI MÃ

(21) 1-2022-02549

(57) Sáng chế đề cập đến phương pháp giải mã, thiết bị giải mã và phương tiện lưu trữ bất biến. Phương pháp bao gồm các bước nhận dòng bit bao gồm tập thông số chuỗi (SPS) chứa cờ thứ nhất và bộ chỉ báo có độ dài của giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC) khi thể hiện trong tiêu đề tham chiếu đến SPS, trong đó giá trị của cờ thứ nhất chỉ định rằng cờ thứ hai có mặt trong tiêu đề tham chiếu đến SPS, và trong đó giá trị của cờ thứ hai chỉ định rằng giá trị POC MSB có mặt trong tiêu đề; xác định giá trị POC dựa trên giá trị POC MSB; xác định ảnh từ dòng bit dựa trên giá trị POC; và giải mã ảnh để thu được ảnh đã được giải mã.

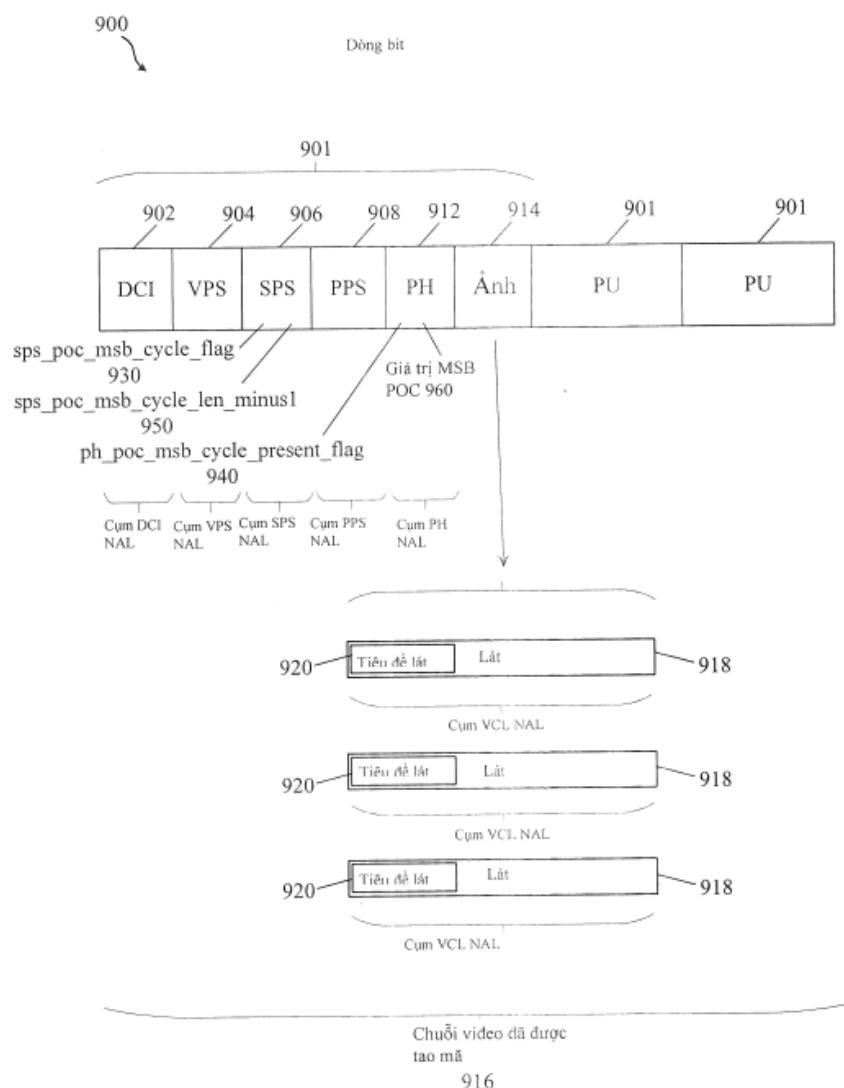


FIG. 9

Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế có liên quan đến các dòng bit video đa lớp trong hoạt động tạo mã video. Cụ thể hơn là, sáng chế có liên quan đến việc hỗ trợ các dòng bit video đa lớp chứa cả ảnh có điểm truy nhập ngẫu nhiên trong ảnh (IRAP: intra random access point) và ảnh không có IRAP.

Tình trạng kỹ thuật của sáng chế

Lượng dữ liệu video được yêu cầu để minh họa thậm chí video tương đối ngắn có thể chắc chắn xảy ra, vốn có thể tạo ra các khó khăn khi dữ liệu đang được phát hoặc truyền thông theo cách khác qua mạng truyền thông với công suất băng thông bị giới hạn. Do vậy, nói chung dữ liệu video được nén trước khi được truyền thông qua các mạng viễn thông hiện đại ngày nay. Kích cỡ của video cũng có thể là vấn đề khi video 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 giữ, từ đó giảm số lượng của dữ liệu cần để biểu diễn các hình ảnh video kỹ thuật số. Sau đó dữ liệu nén được nhận ở đích bởi thiết bị giải nén video để giải mã dữ liệu video. Với các tài nguyên mạng giới hạn và các nhu cầu ngày càng tăng về chất lượng video cao hơn, các công nghệ nén và giải nén được cải thiện mà cải thiện tỷ lệ nén với từ ít tới không bị mất mát chất lượng hình ảnh là điều mong muốn.

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất đề cập đến phương pháp giải mã được thực hiện bởi bộ giải mã video, bao gồm các bước nhận, bởi bộ giải mã video, dòng bit bao gồm tập thông số chuỗi (SPS) và tiêu đề tham chiếu đến SPS, trong đó SPS chứa cờ thứ nhất, trong đó tiêu đề chứa cờ thứ hai, bộ chỉ báo có độ dài của giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC), và giá trị POC MSB, trong đó giá trị của cờ thứ nhất chỉ định rằng cờ thứ hai có mặt trong tiêu đề, và trong đó giá trị của cờ thứ hai chỉ định rằng giá trị POC MSB có mặt trong tiêu đề; xác định, bởi bộ giải mã video, giá trị

POC dựa trên giá trị POC MSB; xác định, bởi bộ giải mã video, ảnh từ dòng bit dựa trên giá trị POC; và giải mã, bởi bộ giải mã video, ảnh để thu được ảnh đã được giải mã.

Phương pháp cung cấp các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất (MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ nhất bao gồm cờ POC MSB được chỉ định là `sps_poc_msb_cycle_flag`.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng giá trị của cờ thứ nhất là một.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ hai bao gồm cờ POC MSB khác và được chỉ định là `ph_poc_msb_cycle_present_flag`.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng giá trị của cờ thứ hai là một.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng giá trị POC được xác định bằng cách móc nối giá trị POC MSB vào giá trị các bit ít đáng kể nhất (LSB) POC.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ hai là cờ mức ảnh.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng bộ chỉ báo của chiều dài cộng 1 chỉ định chiều dài, trong các bit, của POC MSB.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng việc hiển thị ảnh như được giải mã trên màn hiển thị của thiết bị điện tử.

Khía cạnh thứ hai đề cập đến phương pháp mã hóa được thực hiện bởi bộ mã hóa video, phương pháp này bao gồm các bước thiết lập, bởi bộ mã hóa video, cờ thứ nhất trong tập thông số chuỗi (SPS) tới giá trị để chỉ định rằng cờ thứ hai có mặt trong tiêu đề tham chiếu đến SPS; cung cấp trong SPS, bởi bộ mã hóa video, bộ chỉ báo có độ dài của giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC) trong tiêu đề; thiết lập, bởi bộ mã hóa video, cờ thứ hai trong tiêu đề tới giá trị để chỉ định rằng giá trị POC MSB có mặt trong tiêu đề; mã hóa, bởi bộ mã hóa video, SPS và tiêu đề vào trong dòng bit; và lưu trữ, bởi bộ mã hóa video, dòng bit để truyền thông về phía bộ giải mã video.

Phương pháp cung cấp các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất (MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của

các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ nhất bao gồm cờ POC MSB được chỉ định là `sps_poc_msb_cycle_flag`, và trong đó giá trị của cờ thứ nhất là một.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ hai bao gồm cờ POC MSB khác và được chỉ định là `ph_poc_msb_cycle_present_flag`, và trong đó giá trị của cờ thứ hai là một.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng giá trị POC được xác định bằng cách bổ sung giá trị POC MSB vào giá trị các bit ít đáng kể nhất (LSB) POC, và trong đó cờ thứ hai là cờ mức ảnh.

Khía cạnh thứ ba đề cập đến thiết bị giải mã, bao gồm bộ nhận được tạo cấu hình để nhận dòng bit bao gồm tập thông số chuỗi (SPS) chứa cờ thứ nhất và bộ chỉ báo có độ dài của giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC) khi thể hiện trong tiêu đề tham chiếu đến SPS; bộ nhớ được ghép nối với bộ nhận, bộ nhớ để lưu trữ các lệnh; và bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực thi các lệnh khiến cho thiết bị giải mã t0 xác định xem cờ thứ hai có mặt trong tiêu đề tham chiếu đến SPS dựa trên giá trị của cờ thứ nhất; xác định xem giá trị POC MSB có mặt trong tiêu đề dựa trên giá trị của cờ thứ hai khi cờ thứ hai có mặt; xác định giá trị POC, dựa trên giá trị POC MSB khi giá trị POC MSB có mặt; xác định ảnh từ dòng bit dựa trên giá trị POC; và giải mã ảnh để thu được ảnh đã được giải mã

Thiết bị giải mã cung cấp các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất (MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác

nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ nhất bao gồm cờ POC MSB được chỉ định là `sps_poc_msb_cycle_flag`.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ hai bao gồm cờ POC MSB khác và được chỉ định là `ph_poc_msb_cycle_present_flag`.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ hai là cờ mức ảnh.

Khía cạnh thứ tư đề cập đến thiết bị mã hóa, bao gồm bộ nhớ chứa các lệnh; bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh khiến cho thiết bị mã hóa để thiết lập cờ thứ nhất trong tập thông số chuỗi (SPS) tới giá trị để chỉ định rằng cờ thứ hai có mặt trong tiêu đề tham chiếu đến SPS; cung cấp bộ chỉ báo có độ dài của giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC) trong SPS khi thể hiện trong tiêu đề tham chiếu đến SPS; thiết lập cờ thứ hai tới giá trị để chỉ định rằng giá trị POC MSB có mặt trong tiêu đề; và mã hóa SPS và tiêu đề vào trong dòng bit; và bộ truyền được ghép nối với bộ xử lý, bộ truyền được cấu hình để truyền dòng bit về phía bộ giải mã video.

Thiết bị mã hóa cung cấp các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất

(MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ nhất bao gồm cờ POC MSB được chỉ định là `sps_poc_msb_cycle_flag`, và trong đó cờ thứ hai bao gồm cờ POC MSB khác và được chỉ định là `ph_poc_msb_cycle_present_flag`.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng cờ thứ hai là cờ mức ảnh.

Khía cạnh thứ năm đề cập đến thiết bị tạo mã. Thiết bị tạo mã bao gồm bộ nhận được tạo cấu hình để nhận ảnh để mã hóa hoặc để nhận dòng bit để giải mã; bộ truyền được ghép nối với bộ nhận, bộ truyền được cấu hình để truyền dòng bit tới bộ giải mã hoặc để truyền hình ảnh được giải mã tới màn hiển thị; bộ nhớ được ghép nối với ít nhất một trong số bộ nhận hoặc bộ truyền, bộ nhớ được tạo cấu hình để lưu trữ các lệnh; và bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực thi các lệnh được lưu trữ trong bộ nhớ để thực hiện phương pháp bất kỳ trong số các phương pháp được bộc lộ ở đây.

Thiết bị tạo mã cung cấp các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất (MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp

trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh đề xuất rằng màn hiển thị được tạo cấu hình để hiển thị ảnh được giải mã.

Khía cạnh thứ sáu đề cập đến hệ thống. Hệ thống bao gồm bộ mã hóa; và bộ giải mã kết nối với bộ mã hóa, trong đó bộ mã hóa hoặc bộ giải mã bao gồm thiết bị giải mã, thiết bị mã hóa, hoặc thiết bị tạo mã được bọc lộ ở đây.

Hệ thống cung cấp các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất (MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo

mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Khía cạnh thứ bảy đề cập đến phương tiện để tạo mã. Phương tiện để tạo mã bao gồm phương tiện nhận được tạo cấu hình để nhận ảnh để mã hóa hoặc để nhận dòng bit để giải mã; phương tiện truyền được ghép nối với phương tiện nhận, phương tiện truyền được cấu hình để truyền dòng bit tới phương tiện giải mã hoặc để truyền hình ảnh được giải mã tới phương tiện hiển thị; phương tiện lưu trữ được ghép nối với ít nhất một trong số phương tiện nhận hoặc phương tiện truyền, phương tiện lưu trữ được tạo cấu hình để lưu trữ các lệnh; và phương tiện xử lý được ghép nối với phương tiện lưu trữ, phương tiện xử lý được tạo cấu hình để thực thi các lệnh được lưu trữ trong phương tiện lưu trữ để thực hiện phương pháp bất kỳ trong số các phương pháp được bộc lộ ở đây.

Phương tiện để tạo mã cung cấp các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất (MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Nhằm mục đích làm rõ, bất kỳ một trong số các phương án nêu trên có thể được kết hợp với bất kỳ một hoặc nhiều trong số các phương án nêu trên khác để tạo ra phương án mới nằm trong phạm vi của sáng chế.

Các đặc điểm này và các đặc điểm khác sẽ được hiểu rõ hơn từ phần mô tả chi tiết dưới đây kết hợp với các hình vẽ và phần yêu cầu bảo hộ kèm theo.

Mô tả vắn tắt các hình vẽ

Để hiểu hoàn toàn sáng chế, dưới đây tham chiếu đến phần mô tả vắn tắt sau đây, kết hợp với các hình vẽ kèm theo và phần mô tả chi tiết, trong đó các số chỉ dẫn giống nhau thể hiện các phần giống nhau.

FIG.1 là lưu đồ của phương pháp mã hóa tín hiệu video làm ví dụ.

FIG.2 là hình vẽ dạng sơ đồ của hệ thống mã hóa và giải mã (codec) làm ví dụ để mã hóa video.

FIG.3 là hình vẽ dạng giản lược thể hiện bộ mã hóa video làm ví dụ.

FIG.4 là hình vẽ dạng giản lược thể hiện bộ giải mã video làm ví dụ.

FIG.5 minh họa một ví dụ về tạo mã đa lớp đối với khả năng thay đổi không gian.

FIG.6 là hình vẽ thể hiện quan hệ giữa ảnh IRAP tương đối với các ảnh phía trước và các ảnh phía sau trong thứ tự giải mã và thứ tự biểu diễn.

FIG.7 minh họa dòng bit video được tạo cấu hình để thực hiện kỹ thuật làm mới việc giải mã (GDR) dần dần.

FIG.8 là hình vẽ dạng giản lược thể hiện việc tìm kiếm chuyển động không mong muốn khi sử dụng sự hạn chế bộ mã hóa để hỗ trợ GDR.

FIG.9 minh họa một phương án thực hiện của dòng bit video.

FIG.10 thể hiện một phương án thực hiện của phương pháp giải mã dòng bit video được tạo mã.

FIG.11 thể hiện một phương án thực hiện của phương pháp mã hóa dòng bit video được tạo mã.

FIG.12 là hình vẽ dạng giản lược của thiết bị mã hóa video.

FIG.13 là hình vẽ dạng giản lược của một phương án thực hiện của phương tiện để tạo mã.

Mô tả chi tiết sáng chế

Khi bắt đầu cần được hiểu rằng mặc dù cách thực hiện minh họa của một hoặc nhiều phương án được đưa ra dưới đây, các hệ thống và/hoặc các phương pháp đã được bộc lộ có thể được thực hiện sử dụng các kỹ thuật, các kỹ thuật này là đã biết hay hiện có. Sáng chế sẽ không bị giới hạn ở các cách thực hiện minh họa, các hình vẽ, và các kỹ thuật được thể hiện dưới đây, bao gồm các thiết kế mẫu và các cách thực hiện được minh họa và được mô tả dưới đây, mà có thể được cải biến nằm trong phạm vi của yêu cầu bảo hộ kèm theo cùng với phạm vi đầy đủ các phần tương đương của các điểm yêu cầu bảo hộ này.

Sau đây các thuật ngữ được xác định như sau trừ khi được sử dụng trong ngữ cảnh ngược với ngữ cảnh ở đây. Cụ thể là, các định nghĩa sau đây được dự định để cung cấp độ rõ ràng bổ sung cho sáng chế. Tuy nhiên, các thuật ngữ có thể được mô tả theo cách khác nhau trong các ngữ cảnh khác nhau. Theo đó, các định nghĩa sau đây sẽ được xem xét là sự bổ sung và không nên được coi là để giới hạn các định nghĩa khác bất kỳ của các mô tả được tạo ra cho các thuật ngữ đó trong bản mô tả này.

Dòng bit là chuỗi của các bit bao gồm dữ liệu video được nén để truyền giữa bộ mã hóa và bộ giải mã. Bộ mã hóa là thiết bị mà được tạo cấu hình để sử dụng các quy trình mã hóa để nén dữ liệu video vào trong dòng bit. Bộ giải mã là thiết bị mà được tạo cấu hình để sử dụng các quy trình giải mã để tái cấu trúc dữ liệu video từ dòng bit cho việc hiển thị. Ảnh là mảng của các mẫu luma và/hoặc mảng của các mẫu chroma mà tạo ra khung hoặc trường của chúng. Ảnh mà đang được mã hóa hoặc được giải mã có thể được gọi là ảnh hiện tại để làm rõ phần mô tả. Ảnh tham chiếu là ảnh mà chứa các mẫu tham chiếu mà có thể được sử dụng khi tạo mã các ảnh khác nhờ tham chiếu theo dự đoán liên ảnh và/hoặc dự đoán liên lớp. Danh sách ảnh tham chiếu là danh sách của các ảnh tham chiếu dùng cho dự đoán liên ảnh và/hoặc dự đoán liên lớp. Một số hệ thống mã hóa video sử dụng hai danh sách ảnh tham chiếu, mà có thể được biểu thị là danh sách ảnh tham chiếu một và danh sách ảnh tham chiếu không. Cấu trúc danh sách ảnh tham chiếu là cấu trúc cú pháp có thể định địa chỉ mà chứa nhiều danh sách ảnh tham chiếu. Dự đoán liên ảnh là cơ chế tạo mã các mẫu của ảnh hiện tại nhờ tham chiếu đến các mẫu được chỉ báo trong ảnh tham chiếu vốn khác với ảnh hiện tại trong đó ảnh tham chiếu và ảnh hiện tại nằm trong cùng một lớp. Mục

nhập cấu trúc danh sách ảnh tham chiếu là vị trí có thể định địa chỉ trong cấu trúc danh sách ảnh tham chiếu để chỉ báo ảnh tham chiếu được liên kết với danh sách ảnh tham chiếu. Lát cắt tiêu đề là một phần của lát cắt được tạo mã chứa các phần tử dữ liệu thuộc về tất cả dữ liệu video bên trong ảnh xếp lớp được biểu diễn trong lát cắt. Tập thông số chuỗi (SPS) là tập thông số mà chứa dữ liệu liên quan đến chuỗi của các ảnh. Tập thông số ảnh (PPS) là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều ảnh toàn cục được tạo mã như được xác định bởi thành phần cú pháp được tìm thấy trong mỗi tiêu đề ảnh.

Cờ là thành phần cú pháp biến thiên hoặc một bit đơn mà có thể mang một trong số hai trị số khả dĩ: 0 và 1. Cụm truy nhập (AU: access unit) là tập hợp của một hoặc nhiều ảnh được tạo mã được kết hợp với cùng thời gian hiển thị (ví dụ, số đếm thứ tự ảnh giống nhau) cho việc xuất ra từ bộ đệm hình ảnh đã được giải mã (DPB: decoded picture buffer) (ví dụ, để hiển thị cho người dùng). Bộ tách cụm truy nhập (AUD) là bộ chỉ báo hoặc cấu trúc dữ liệu được sử dụng để chỉ báo sự bắt đầu của AU hoặc đường bao giữa các AU. Chuỗi video được giải mã là chuỗi của các ảnh mà đã được tái tạo bởi bộ giải mã khi chuẩn bị để hiển thị cho người dùng.

Chuỗi video đã được tạo mã (CVS: coded video sequence) là chuỗi của các cụm truy nhập (AU) mà bao gồm, theo thứ tự giải mã, AU bắt đầu chuỗi video đã được tạo mã (CVSS: coded video sequence start), theo sau bởi không hoặc nhiều AU mà không là CVSS AU, bao gồm tất cả các AU tiếp theo lên tới nhưng không bao gồm AU tiếp theo bất kỳ mà là CVSS AU. CVSS AU là AU trong đó có cụm dự đoán (PU: prediction unit) cho mỗi lớp được định rõ bởi tập thông số video (VPS) và ảnh đã được tạo mã trong mỗi PU là ảnh bắt đầu chuỗi video lớp được tạo mã (CLVSS). Theo một phương án, mỗi ảnh nằm bên trong AU. PU 901 là tập hợp của các cụm lớp trừu tượng mạng (NAL: network abstraction layer) mà được liên kết với nhau theo quy tắc phân loại được định rõ, là liên tiếp theo thứ tự giải mã, và chứa chính xác một ảnh đã được tạo mã.

Số đếm thứ tự ảnh (POC) là biến mà được liên kết với mỗi ảnh, chỉ xác định ảnh kết hợp trong số các tất cả các ảnh trong CLVS, và, khi ảnh kết hợp sẽ được xuất ra từ DPB, chỉ báo vị trí của ảnh kết hợp theo thứ tự xuất tương đối với các vị trí thứ tự đầu ra của các ảnh khác trong CLVS giống nhau mà sẽ được xuất ra từ DPB. Bit có

nghĩa tối đa (MSB, cũng gọi là bit thứ tự cao) là vị trí bit trong số nhị phân có giá trị lớn nhất. MSB đôi khi được gọi là bit thứ tự cao hoặc bit tận cùng bên trái do quy ước về ký hiệu vị trí khi viết các con số đáng kể hơn lệch về bên trái. Tiêu đề ảnh (PH) là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho tất cả các lát cắt của ảnh đã được tạo mã.

Các từ viết tắt sau đây được sử dụng trong bản mô tả: cụm truy nhập (AU), cây tạo mã khối (CTB), cụm cây tạo mã (CTU: Coding Tree Unit), cụm tạo mã (CU: Coding Unit), chuỗi video lớp được tạo mã (CLVS), bắt đầu chuỗi video lớp được tạo mã (CLVSS), chuỗi video đã được tạo mã (CVS: coded video sequence), bắt đầu chuỗi video đã được tạo mã (CVSS), nhóm chuyên gia video liên kết (JVET: Joint Video Experts Team), tập hợp ảnh xếp lớp được hạn chế chuyển động (MCTS), cụm truyền tối đa (MTU: Maximum Transfer Unit), lớp trừu tượng mạng (NAL: network abstraction layer), tập hợp lớp đầu ra (OLS), số đếm thứ tự ảnh (POC), tập thông số ảnh (PPS), điểm truy nhập ngẫu nhiên (RAP), phần tải dữ liệu chuỗi byte thô (RBSP: Raw Byte Sequence Payload), tập thông số chuỗi (SPS), tập thông số video (VPS), tạo mã video vạn năng (VVC), và bản thảo công tác (WD: Working Draft).

FIG.1 là lưu đồ của phương pháp hoạt động làm ví dụ 100 để tạo mã tín hiệu video. Cụ thể là, tín hiệu video được mã hóa ở bộ mã hóa. Quá trình mã hóa nén tín hiệu video bằng cách sử dụng các cơ chế khác nhau để làm giảm tệp kích cỡ video. Kích cỡ tệp nhỏ hơn cho phép tệp video nén được truyền về phía người dùng, đồng thời làm giảm chi phí băng thông liên quan. Bộ giải mã sau đó giải mã tệp video nén để tái cấu trúc tín hiệu video ban đầu để hiển thị cho người dùng đầu cuối. Quá trình giải mã nói chung ngược với quá trình mã hóa để cho phép bộ giải mã tái cấu trúc một cách đồng nhất tín hiệu video.

Ở bước 101, tín hiệu video được nhập vào trong bộ mã hóa. Ví dụ, tín hiệu video có thể là tệp video không nén chứa trong bộ nhớ. Theo một ví dụ khác, tệp video có thể được ghi bởi thiết bị ghi video, như máy quay video, và được mã hóa để hỗ trợ phát video trực tiếp. Tệp video có thể bao gồm cả thành phần âm thanh và thành phần video. Thành phần video chứa một dãy các khung hình ảnh mà, khi được xem liên tục, tạo ra sự ấn tượng bằng mắt của chuyển động. Các khung chứa các điểm ảnh mà được thể hiện là do ánh sáng, được nói đến ở đây như các thành phần luma (hoặc các mẫu

luma), và màu sắc, mà được gọi là các thành phần chroma (hoặc các mẫu màu sắc). Theo một số ví dụ, các khung cũng có thể chứa các giá trị chiều sâu để hỗ trợ chế độ xem ba chiều.

Ở bước 103, video được phân chia thành các khối. Việc phân chia bao gồm chia nhỏ các điểm ảnh trong mỗi khung thành các khối hình vuông và/hoặc hình chữ nhật cho việc nén. Ví dụ, trong tạo mã video hiệu suất cao (HEVC: High Efficiency Video Coding) (còn được biết đến là H.265 và MPEG-H Phần 2), khung trước tiên có thể được chia thành các cụm cây tạo mã (CTUs), vốn là các khối có kích cỡ định trước (ví dụ, 64 điểm ảnh x 64 điểm ảnh). Các CTU chứa cả mẫu luma và chroma. Các cây mã hóa có thể được sử dụng để chia các CTU thành các khối và sau đó chia nhỏ đệ quy các khối cho đến khi các kết cấu đạt được để tiếp tục hỗ trợ mã hóa. Ví dụ, các thành phần luma của khung có thể được chia nhỏ cho đến khi các khối riêng biệt chứa các giá trị ánh sáng tương đối đồng nhất. Ngoài ra, các thành phần chroma của khung có thể được chia nhỏ cho đến khi các khối riêng biệt chứa các giá trị màu sắc tương đối đồng nhất. Theo đó, các cơ chế phân chia thay đổi tùy thuộc vào thành phần của các khung video.

Ở bước 105, khác nhau các cơ chế nén được sử dụng để nén các khối hình ảnh mà được phân chia ở bước 103. Ví dụ, sự dự đoán ngoại suy và/hoặc dự đoán nội suy có thể được sử dụng. Dự đoán ngoại suy được thiết kế để thu được ưu điểm theo thực tế là các đối tượng trong trường hợp chung có xu hướng xuất hiện trong các khung liên tiếp. Theo đó, khối thể hiện đối tượng trong khung tham chiếu không cần được mô tả lặp lại ở các khung liên kế. Cụ thể là, đối tượng, như bảng, có thể vẫn ở vị trí không đổi trên nhiều khung. Do đó bảng được mô tả một lần và các khung liên kế có thể lại tham chiếu đến khung tham chiếu. Các cơ chế đối chiếu mẫu hình có thể được sử dụng để so khớp với các đối tượng trên nhiều khung. Ngoài ra, các đối tượng chuyển động có thể được biểu diễn qua nhiều khung, ví dụ do chuyển động đối tượng hoặc chuyển động máy ghi hình. Theo ví dụ cụ thể, video có thể thể hiện ô tô mà dịch chuyển qua màn hình trên nhiều khung. Các vector chuyển động có thể được sử dụng để mô tả chuyển động này. Vector chuyển động là vector hai chiều mà tạo ra độ lệch từ các tọa độ của đối tượng trong khung đến các tọa độ của đối tượng trong khung tham chiếu. Như vậy, dự đoán ngoại suy có thể mã hóa khối hình ảnh trong khung hiện tại thành

tập hợp của các vectơ chuyển động biểu thị độ lệch từ khối tương ứng trong khung tham chiếu.

Dự đoán nội suy mã hóa các khối trong khung chung. Dự đoán nội suy mang ưu điểm là các thành phần luma và chroma có xu hướng tụ lại trong khung. Ví dụ, mảng màu xanh lục trong một phần của cây có xu hướng được định vị liền kề với các mảng màu xanh lục tương tự. Dự đoán nội suy sử dụng nhiều chế độ dự đoán có hướng (ví dụ, ba mươi ba trong HEVC), chế độ phẳng, và chế độ dòng điện một chiều (DC). Các chế độ có hướng chỉ báo mà khối hiện tại là tương tự/giống với mẫu của khối lân cận theo hướng tương ứng. Chế độ phẳng chỉ báo rằng một dãy các khối dọc theo hàng/cột (ví dụ, mặt phẳng) có thể được nội suy dựa trên các khối lân cận ở các mép của hàng. Chế độ phẳng, trong tác dụng, biểu thị sự chuyển tiếp trơn nhẹ của ánh sáng/màu sắc qua hàng/cột bằng cách sử dụng độ dốc tương đối không đổi trong các giá trị thay đổi. Chế độ DC được sử dụng cho làm mịn đường biên và chỉ báo rằng khối là tương tự/giống với trị số trung bình kết hợp với mẫu của tất cả các khối lân cận kết hợp với các hướng có góc của các chế độ dự đoán có hướng. Theo đó, các khối dự đoán nội suy có thể thể hiện các khối hình ảnh dưới dạng các giá trị chế độ dự đoán tương đối khác nhau thay vì các giá trị thực. Ngoài ra, các khối dự đoán ngoại suy có thể thể hiện các khối hình ảnh dưới dạng các giá trị vectơ chuyển động thay vì các giá trị thực. Trong cả hai trường hợp, các khối dự đoán không thể thể hiện một cách chính xác các khối hình ảnh trong một số trường hợp. Các khác biệt bất kỳ được chứa trong các khối dư. Các biến đổi có thể được áp dụng cho các khối dư để tiếp tục nén tệp tin.

Ở bước 107, các công nghệ lọc khác nhau có thể được áp dụng. Trong HEVC, các bộ lọc được áp dụng theo sơ đồ lọc khép kín. Sự dự đoán dựa trên khối nêu trên có thể dẫn tới sự hình thành của các khối hình ảnh ở bộ giải mã. Ngoài ra, sự dự đoán dựa trên sơ đồ khối có thể mã hóa khối và sau đó tái cấu trúc khối được mã hóa để sử dụng về sau làm khối tham chiếu. Sơ đồ lọc khép kín áp dụng lặp lại các bộ lọc nén nhiều, các bộ lọc giải khối, các bộ lọc vòng lặp thích ứng, và các bộ lọc độ lệch thích ứng mẫu (SAO: sample adaptive offset) cho các khối/các khung. Các bộ lọc này làm giảm các thành phần khối lạ như vậy sao cho tệp tin được mã hóa có thể được tái cấu trúc một cách chính xác. Ngoài ra, các bộ lọc này làm giảm các thành phần lạ trong các khối tham chiếu tái cấu trúc sao cho các thành phần lạ hầu như ít tạo ra các thành phần

khác lạ trong các khối tiếp sau mà được mã hóa dựa trên các khối tham chiếu tái cấu trúc.

Khi tín hiệu video được phân chia, nén, và lọc, dữ liệu tạo thành được mã hóa trong dòng bit ở bước 109. Dòng bit bao gồm dữ liệu nêu trên cũng như dữ liệu truyền tín hiệu bất kỳ được kỳ vọng để hỗ trợ thích hợp sự tái cấu trúc tín hiệu video ở bộ giải mã. Ví dụ, dữ liệu này có thể bao gồm dữ liệu phân chia, dữ liệu dự đoán, các khối dư, và các cờ khác nhau cung cấp các lệnh mã hóa cho bộ giải mã. Dòng bit có thể có mặt trong bộ nhớ để truyền về phía bộ giải mã ngay khi yêu cầu. Dòng bit cũng có thể là phát rộng và/hoặc phát đa phương về phía các bộ giải mã. Sự hình thành của dòng bit là quá trình lặp. Theo đó, các bước 101, 103, 105, 107, và 109 có thể xảy ra một cách liên tục và/hoặc một cách đồng thời trên nhiều khung và nhiều khối. Thứ tự được thể hiện trên FIG.1 được thể hiện để giải thích rõ và làm đơn giản phần mô tả, và không được dự tính để giới hạn việc xử lý mã hóa video ở thứ tự cụ thể.

Bộ giải mã nhận dòng bit và bắt đầu quá trình giải mã ở bước 111. Cụ thể là, bộ giải mã sử dụng sơ đồ giải mã entropy để chuyển đổi dòng bit thành cú pháp tương ứng và dữ liệu video. Bộ giải mã sử dụng dữ liệu cú pháp từ dòng bit để xác định các phân chia cho các khung ở bước 111. Việc phân chia nên phù hợp với các kết quả của phân chia khối ở bước 103. Việc mã hóa/giải mã entropy như được sử dụng ở bước 111 được mô tả dưới đây. Bộ mã hóa tạo nhiều lựa chọn trong quá trình nén, như chọn các sơ đồ phân chia khối từ một số lựa chọn có thể có dựa trên định vị không gian của các giá trị trong (các) hình ảnh đầu vào. Truyền tín hiệu các lựa chọn chính xác có thể sử dụng số lượng lớn bin. Như được sử dụng ở đây, bin là trị số nhị phân mà được coi là biến (ví dụ, trị số bit có thể khác nhau tùy thuộc vào ngữ cảnh). Mã entropy cho phép bộ mã hóa loại bỏ các tùy chọn bất kỳ mà không thể tồn tại một cách rõ ràng cho trường hợp cụ thể, để lại tập hợp của các tùy chọn có thể được phép. Mỗi tùy chọn có thể được phép sau đó được gán từ mã. Chiều dài của các từ mã được dựa trên số lượng các tùy chọn cho phép (ví dụ, một bin cho hai tùy chọn, hai bin cho ba đến bốn tùy chọn, v.v.) Sau đó bộ mã hóa mã hóa từ mã cho tùy chọn được chọn. Sơ đồ này làm giảm kích cỡ của các từ mã khi các từ mã ngày càng lớn theo mong muốn để chỉ báo duy nhất sự lựa chọn từ các tập con nhỏ của các tùy chọn có thể được phép do trái ngược với việc biểu thị duy nhất sự lựa chọn từ tập có thể là lớn nhất của tất cả các tùy

chọn có thể có. Bộ giải mã sau đó giải mã sự lựa chọn bằng cách xác định tập hợp của các tùy chọn có thể được phép theo cách tương tự với bộ mã hóa. Bằng cách xác định tập hợp của các tùy chọn có thể được phép, bộ giải mã có thể đọc từ mã và xác định sự lựa chọn được tạo bởi bộ mã hóa.

Ở bước 113, bộ giải mã thực hiện giải mã khối. Cụ thể là, bộ giải mã sử dụng các biến đổi ngược để tạo ra các khối dư. Sau đó bộ giải mã sử dụng các khối dư và các khối dự đoán tương ứng để tái cấu trúc các khối hình ảnh theo việc phân chia. Các khối dự đoán có thể bao gồm cả các khối dự đoán nội suy và các khối dự đoán ngoại suy như được tạo ra ở bộ mã hóa ở bước 105. Các khối hình ảnh được tái cấu trúc sau đó được định vị vào trong các khung của tín hiệu video tái cấu trúc theo việc phân chia dữ liệu được xác định ở bước 111. Cú pháp cho bước 113 cũng có thể được tạo tín hiệu trong dòng bit thông qua mã entropy như đề cập đến bên trên.

Ở bước 115, việc lọc được thực hiện trên các khung của tín hiệu video tái cấu trúc theo cách tương tự với bước 107 ở bộ mã hóa. Ví dụ, các bộ lọc nén nhiễu, các bộ lọc giải khối, các bộ lọc vòng lặp thích ứng, và SAO các bộ lọc có thể được áp dụng cho các khung để loại bỏ các thành phần lạ của khối. Khi các khung được lọc, tín hiệu video có thể được gửi ra đến màn hiển thị ở bước 117 để xem bởi người dùng đầu cuối.

FIG.2 là hình vẽ dạng sơ đồ của hệ thống mã hóa và giải mã (codec) làm ví dụ 200 để mã hóa video. Cụ thể là, hệ thống mã hóa 200 cung cấp chức năng để hỗ trợ việc thực hiện của phương pháp hoạt động 100. Hệ thống mã hóa 200 được tạo ra để minh họa các thành phần được sử dụng ở cả bộ mã hóa và bộ giải mã. Hệ thống mã hóa 200 nhận và phân chia tín hiệu video như được mô tả đối với các bước 101 và 103 trong phương pháp hoạt động 100, mà tạo ra tín hiệu video đã phân chia 201. Hệ thống mã hóa 200 sau đó nén tín hiệu video đã phân chia 201 thành dòng bit mã hóa khi hoạt động như bộ mã hóa như được mô tả đối với các bước 105, 107, và 109 trong phương pháp 100. Khi hoạt động như hệ thống bộ giải mã mã hóa 200 sinh ra đầu ra tín hiệu video từ dòng bit như được mô tả đối với các bước 111, 113, 115, và 117 trong phương pháp hoạt động 100. Hệ thống mã hóa 200 bao gồm thành phần điều khiển bộ mã hóa chung 211, thành phần chuyển đổi tỷ lệ và lượng tử hóa 213, thành phần ước tính bên trong ảnh 215, thành phần dự đoán nội suy ảnh 217, thành phần bù chuyển

động 219, thành phần ước tính chuyển động 221, thành phần chuyển đổi tỷ lệ và nghịch đảo 229, thành phần phân tích điều khiển bộ lọc 227, thành phần bộ lọc vòng kín 225, thành phần đệm hình ảnh giải mã 223, và thành phần định dạng tiêu đề và mã hóa số học nhị phân tương thích ngữ cảnh (CABAC) 231. Các thành phần này được ghép nối như được thể hiện. Trên FIG.2, các đường màu đen chỉ báo chuyển động của dữ liệu sẽ được mã hóa/giải mã mặc dù các đường nét đứt chỉ báo chuyển động của dữ liệu điều khiển để điều khiển sự hoạt động của các bộ phận khác. Các thành phần của hệ thống mã hóa 200 có thể có mặt tất cả trong bộ mã hóa. Bộ giải mã có thể bao gồm tập hợp con của các thành phần của hệ thống mã hóa 200. Ví dụ, bộ giải mã có thể bao gồm thành phần dự đoán nội suy ảnh 217, thành phần bù chuyển động 219, thành phần chuyển đổi tỷ lệ và nghịch đảo 229, thành phần bộ lọc vòng kín 225, và thành phần đệm hình ảnh giải mã 223. Các bộ phận này được mô tả dưới đây.

Tín hiệu video đã phân chia 201 là chuỗi video được ghi mà đã được phân chia thành các khối các điểm ảnh bởi cây mã hóa. Cây mã hóa sử dụng các chế độ phân chia khác nhau để chia nhỏ khối các điểm ảnh thành các khối nhỏ hơn của các điểm ảnh. Sau đó các khối này có thể tiếp tục được chia nhỏ thành các khối nhỏ hơn. Các khối có thể được gọi là các nút trên cây mã hóa. Các nút cha lớn hơn được phân chia thành các nút con nhỏ hơn. Số lần mà nút được chia nhỏ được gọi là chiều sâu của nút/cây mã hóa. Các khối đã chia có thể có mặt trong cụm tạo mã (CU: coding unit) trong một số trường hợp. Ví dụ, CU có thể là phần con của CTU chứa khối luma, (các) khối chroma (Cr) chênh lệch màu đỏ, và (các) khối chroma (Cb) chênh lệch xanh dương) cùng với các lệnh cú pháp tương ứng cho CU. Các chế độ phân chia có thể bao gồm cây nhị phân (BT: binary tree), cây tam phân (TT: triple tree), và cây tứ phân (QT: quad tree) được sử dụng để phân chia nút thành hai, ba, hoặc bốn nút con, một cách tương ứng), có các hình dạng khác nhau tùy thuộc vào các chế độ phân chia được sử dụng. Tín hiệu video đã phân chia 201 được chuyển tiếp đến thành phần điều khiển bộ mã hóa chung 211, thành phần chuyển đổi tỷ lệ và lượng tử hóa 213, thành phần ước tính bên trong ảnh 215, thành phần phân tích điều khiển bộ lọc 227, và thành phần ước tính chuyển động 221 cho việc nén.

Thành phần điều khiển bộ mã hóa chung 211 được cấu hình để tạo các quyết định liên quan đến việc mã hóa các hình ảnh của chuỗi video thành dòng bit các giới

hạn ứng dụng. Ví dụ, thành phần điều khiển bộ mã hóa chung 211 quản lý sự tối ưu hóa của tốc độ bit/kích cỡ dòng bit so với chất lượng tái cấu trúc. Các quyết định này có thể được tạo ra dựa trên khoảng trống chứa/băng thông khả dụng và yêu cầu độ phân giải hình ảnh. Thành phần điều khiển bộ mã hóa chung 211 cũng quản lý Việc sử dụng bộ đệm theo tốc độ truyền để làm giảm các sự cố chạy dưới mức và chạy quá mức bộ đệm. Để quản lý các sự cố này, thành phần điều khiển bộ mã hóa chung 211 quản lý sự phân chia, dự đoán, và việc lọc nhờ các thành phần khác. Ví dụ, thành phần điều khiển bộ mã hóa chung 211 có thể tăng theo dạng động độ phức tạp nén để tăng độ phân giải và tăng việc sử dụng băng thông hoặc giảm độ phức tạp nén để giảm độ phân giải và việc sử dụng băng thông. Do đó, thành phần điều khiển bộ mã hóa chung 211 điều khiển các thành phần khác của hệ thống mã hóa 200 để cân bằng chất lượng tái cấu trúc tín hiệu video với các mối quan tâm về tốc độ bit. Thành phần điều khiển bộ mã hóa chung 211 tạo ra dữ liệu điều khiển, để điều khiển sự hoạt động của các thành phần khác. Dữ liệu điều khiển cũng được chuyển tiếp đến thành phần định dạng tiêu đề và CABAC 231 sẽ được mã hóa trong dòng bit thành các thông số tín hiệu để giải mã ở bộ giải mã.

Tín hiệu video đã phân chia 201 cũng được gửi đến thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 để đoán ngoại suy. Khung hoặc lát cắt của tín hiệu video đã phân chia 201 có thể được chia thành nhiều khối video. Thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 thực hiện mã hóa dự đoán bên trong khối video nhận được tương đối với một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu để tạo ra dự đoán tạm thời. Hệ thống mã hóa 200 có thể thực hiện nhiều lần mã hóa, ví dụ, để chọn chế độ mã hóa thích hợp cho mỗi khối của dữ liệu video.

Thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp ở mức cao, nhưng được minh họa một cách riêng biệt cho các mục đích thuộc về lý thuyết. Việc ước tính chuyển động, được thực hiện bởi thành phần ước tính chuyển động 221, là quá trình tạo ra các vectơ chuyển động, để ước tính chuyển động cho các khối video. Vectơ chuyển động, ví dụ, có thể chỉ báo sự dịch chuyển của đối tượng được mã hóa tương đối với khối dự đoán. Khối dự đoán là khối được tìm ra là khối phù hợp nhất để được mã hóa, nhờ sự chênh lệch điểm ảnh. Khối

dự đoán cũng có thể được gọi là khối tham chiếu. Sự chênh lệch điểm ảnh này có thể được xác định bởi tổng cộng của tuyệt đối độ chênh lệch (SAD), tổng cộng của độ chênh lệch bình phương (SSD), hoặc các số đo độ chênh lệch khác. HEVC sử dụng một số đối tượng mã hóa bao gồm CTU, các khối cây mã hóa (CTB: coding tree block), và các CU. Ví dụ, CTU có thể được chia thành các CTB, mà sau đó có thể được chia thành các CB để có mặt trong các CU. CU có thể được mã hóa thành đơn vị dự đoán (PU) chứa dữ liệu dự đoán và/hoặc đơn vị chuyển đổi (TU) chứa dữ liệu dư đã được chuyển đổi cho CU. Thành phần ước tính chuyển động 221 sinh ra các vector chuyển động, các PU, và các TU nhờ sử dụng phân tích tốc độ méo như một phần của quá trình tối ưu tốc độ méo. Ví dụ, thành phần ước tính chuyển động 221 có thể xác định nhiều khối tham chiếu, nhiều vector chuyển động, v.v. cho khối hiện tại/khung, và có thể chọn các khối tham chiếu, các vector chuyển động, v.v. có các đặc tính tốc độ méo tốt nhất. Các đặc tính tốc độ méo làm cân bằng cả chất lượng của video sự tái cấu trúc (ví dụ, lượng dữ liệu thất thoát do nén) với hiệu suất mã hóa (ví dụ, kích cỡ của sự mã hóa cuối cùng).

Theo một số ví dụ, hệ thống mã hóa 200 có thể tính toán các giá trị cho các vị trí điểm ảnh số nguyên con của ảnh tham chiếu chứa trong thành phần đệm hình ảnh giải mã 223. Ví dụ, hệ thống mã hóa video 200 có thể nội suy các giá trị của vị trí điểm ảnh một phần tư, vị trí điểm ảnh một phần tám, hoặc vị trí điểm ảnh phân số khác của hình ảnh tham chiếu. Do đó, thành phần ước tính chuyển động 221 có thể thực hiện tìm kiếm chuyển động tương đối với vị trí điểm ảnh đầy đủ và vị trí điểm ảnh phân số và gửi ra vector chuyển động với độ chính xác điểm ảnh phân số. Thành phần ước tính chuyển động 221 tính toán vector chuyển động cho PU của khối video trong lát cắt được mã hóa bên trong bằng cách so sánh vị trí của PU với vị trí của khối dự đoán của hình ảnh tham chiếu. Thành phần ước tính chuyển động 221 gửi ra vector chuyển động đã được tính toán dưới dạng dữ liệu chuyển động đến định dạng tiêu đề và thành phần CABAC 231 để mã hóa và dịch chuyển đến thành phần bù chuyển động 219.

Việc bù chuyển động, được thực hiện bởi thành phần bù chuyển động 219, có thể bao gồm sự truy tìm hoặc tạo ra khối dự đoán dựa trên vector chuyển động được xác định bởi thành phần ước tính chuyển động 221. Ngoài ra, thành phần ước tính

chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp theo chức năng, theo một số ví dụ. Ngay khi nhận vector chuyển động cho PU của khối video hiện tại, thành phần bù chuyển động 219 có thể định vị khối dự đoán mà vector chuyển động hướng vào đó. Sau đó khối video dư được tạo bằng cách trừ đi các giá trị điểm ảnh của khối dự đoán từ các giá trị điểm ảnh của khối video hiện tại được mã hóa, tạo ra các giá trị chênh lệch điểm ảnh. Nói chung, thành phần ước tính chuyển động 221 thực hiện ước tính chuyển động tương đối với các thành phần luma, và thành phần bù chuyển động 219 sử dụng các vector chuyển động được tính toán dựa trên các thành phần luma cho cả các thành phần chroma và các thành phần luma. Khối dự đoán và khối dư được chuyển tiếp đến thành phần chuyển đổi tỷ lệ và lượng tử hóa 213.

Tín hiệu video đã phân chia 201 cũng được gửi đến thành phần ước tính bên trong ảnh 215 và thành phần dự đoán nội suy ảnh 217. Như với thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219, thành phần ước tính bên trong ảnh 215 và thành phần dự đoán nội suy ảnh 217 có thể được tích hợp ở mức cao, nhưng được minh họa một cách riêng biệt cho các mục đích thuộc về lý thuyết. Thành phần ước tính bên trong ảnh 215 và thành phần dự đoán nội suy ảnh 217 dự đoán nội suy khối hiện tại tương đối với các khối trong khung hiện tại, như một sự thay thế cho dự đoán ngoại suy được thực hiện bởi thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219 giữa các khung, như được mô tả trên đây. Cụ thể là, thành phần ước tính bên trong ảnh 215 xác định chế độ dự đoán nội suy để sử dụng để mã hóa khối hiện tại. Theo một số ví dụ, thành phần ước tính bên trong ảnh 215 chọn chế độ dự đoán nội suy thích hợp để mã hóa khối hiện tại từ nhiều chế độ dự đoán nội suy đã được thử nghiệm. Các chế độ dự đoán nội suy được chọn sau đó được chuyển tiếp đến bước định dạng tiêu đề và thành phần CABAC 231 để mã hóa.

Ví dụ, thành phần ước tính bên trong ảnh 215 tính toán các giá trị tốc độ méo nhờ sử dụng phân tích tốc độ méo cho các chế độ dự đoán nội suy được thử nghiệm khác nhau, và chọn chế độ dự đoán nội suy có các đặc tính tốc độ méo trong số các chế độ được thử nghiệm. Phân tích tốc độ méo nói chung xác định lượng độ méo (hoặc lỗi) giữa khối được mã hóa và khối không được mã hóa ban đầu mà được mã hóa để tạo ra khối được mã hóa, cũng như tốc độ bit (ví dụ, số lượng bit) được sử dụng để tạo ra khối được mã hóa. Thành phần ước tính bên trong ảnh 215 tính toán các tỷ số

từ các nhiễu loạn và các tốc độ cho các khối được mã hóa khác nhau để xác định chế độ dự đoán nội suy nào có trị số tốc độ méo tốt nhất cho khối. Ngoài ra, thành phần ước tính bên trong ảnh 215 có thể được cấu hình để mã hóa chiều sâu các khối của bản đồ chiều sâu nhờ sử dụng chế độ mô hình hóa chiều sâu (DMM) dựa trên sự tối ưu hóa tốc độ méo (RDO).

Thành phần dự đoán nội suy ảnh 217 có thể tạo ra khối dư từ khối dự đoán dựa trên các chế độ dự đoán nội suy được chọn được xác định bởi thành phần ước tính bên trong ảnh 215 khi được thực hiện trên bộ mã hóa hoặc đọc khối dư từ dòng bit khi được thực hiện trên bộ giải mã. Khối dư bao gồm độ chênh lệch trong các giá trị giữa khối dự đoán và khối ban đầu, được biểu diễn là ma trận. Khối dư sau đó được chuyển tiếp đến thành phần chuyển đổi tỷ lệ và lượng tử hóa 213. Thành phần ước tính bên trong ảnh 215 và thành phần dự đoán nội suy ảnh 217 có thể dẫn động ở cả các thành phần luma và chroma.

Thành phần chuyển đổi tỷ lệ và lượng tử hóa 213 được cấu hình để tiếp tục nén khối dư. Thành phần chuyển đổi tỷ lệ và lượng tử hóa 213 áp dụng sự chuyển đổi, như chuyển đổi cosin rời rạc (DCT), chuyển đổi sin rời rạc (DST), hoặc chuyển đổi tương tự theo lý thuyết, cho khối dư, tạo ra khối video bao gồm các giá trị hệ số chuyển đổi dư. Các biến đổi sóng nhỏ, các biến đổi nguyên, các biến đổi dải con hoặc các loại khác của các biến đổi cũng có thể được sử dụng. Chuyển đổi có thể thông tin chuyển đổi dư từ miền trị số điểm ảnh thành miền chuyển đổi, như miền tần số. Thành phần chuyển đổi tỷ lệ và lượng tử hóa 213 cũng được tạo cấu hình để định tỷ lệ thông tin dư đã được chuyển đổi, ví dụ dựa trên tần số. Việc định tỷ lệ này bao gồm áp dụng các yếu tố tỷ lệ vào thông tin dư sao cho thông tin tần số khác nhau được lượng tử hóa ở các độ chi tiết khác nhau, mà có thể ảnh hưởng đến chất lượng nhìn thấy cuối cùng của video được tái cấu trúc. Thành phần chuyển đổi tỷ lệ và lượng tử hóa 213 cũng được tạo cấu hình để lượng tử hóa các hệ số chuyển đổi để tiếp tục làm giảm tốc độ bit. Quá trình lượng tử hóa có thể làm giảm chiều sâu bit kết hợp với một số hoặc tất cả trong số các hệ số. Độ của lượng tử hóa có thể được thay đổi bằng cách điều chỉnh thông số lượng tử hóa. Theo một số ví dụ, thành phần chuyển đổi tỷ lệ và lượng tử hóa 213 sau đó có thể thực hiện quét của ma trận bao gồm các hệ số chuyển đổi được lượng tử hóa.

Các hệ số chuyển đổi được lượng tử hóa được chuyển tiếp đến bước định dạng tiêu đề và thành phần CABAC 231 sẽ được mã hóa trong dòng bit.

Thành phần chuyển đổi tỷ lệ và nghịch đảo 229 áp dụng sự hoạt động ngược của thành phần chuyển đổi tỷ lệ và lượng tử hóa 213 để hỗ trợ sự ước tính chuyển động. Thành phần chuyển đổi tỷ lệ và nghịch đảo 229 áp dụng tỷ lệ nghịch, sự chuyển đổi, và/hoặc lượng tử hóa để tái cấu trúc khối dư trong miền điểm ảnh, ví dụ, để sử dụng về sau làm khối tham chiếu mà có thể trở thành khối dự đoán cho khối hiện tại khác. Thành phần ước tính chuyển động 221 và/hoặc thành phần bù chuyển động 219 có thể tính toán khối tham chiếu bằng cách bổ sung khối dư trở lại cho khối dự đoán tương ứng để sử dụng trong ước tính chuyển động của khối/khung phía sau. Các bộ lọc được áp dụng cho các khối tham chiếu tái cấu trúc để làm giảm các thành phần lạ được tạo ra trong quá trình định tỷ lệ, lượng tử hóa, và chuyển đổi. Các thành phần lạ này có thể theo cách khác gây ra sự dự đoán không chính xác (và tạo ra các thành phần lạ khác) khi các khối tiếp sau được dự đoán.

Thành phần phân tích điều khiển bộ lọc 227 và thành phần bộ lọc vòng kín 225 áp dụng các bộ lọc cho các khối dư và/hoặc cho các khối hình ảnh được tái cấu trúc. Ví dụ, khối dư đã chuyển đổi từ thành phần chuyển đổi tỷ lệ và nghịch đảo 229 có thể được kết hợp với khối dự đoán tương ứng từ thành phần dự đoán nội suy ảnh 217 và/hoặc thành phần bù chuyển động 219 để tái cấu trúc khối hình ảnh ban đầu. Các bộ lọc sau đó có thể được áp dụng cho khối hình ảnh được tái cấu trúc. Theo một số ví dụ, theo cách thay thế, các bộ lọc có thể được áp dụng cho các khối dư. Như với các bộ phận khác trên FIG.2, thành phần phân tích điều khiển bộ lọc 227 và thành phần bộ lọc vòng kín 225 được tích hợp mức cao và có thể được thực hiện với nhau, nhưng được minh họa một cách riêng biệt cho các mục đích thuộc về lý thuyết. Các bộ lọc mà được áp dụng cho các khối tham chiếu tái cấu trúc được áp dụng cho các vùng không gian cụ thể và bao gồm nhiều thông số để điều chỉnh cách mà các bộ lọc này được áp dụng. Thành phần phân tích điều khiển bộ lọc 227 phân tích các khối tham chiếu tái cấu trúc để xác định nơi mà các bộ lọc này nên được áp dụng và thiết lập các thông số tương ứng. Dữ liệu này được chuyển tiếp đến thành phần định dạng tiêu đề và CABAC 231 dưới dạng dữ liệu điều khiển bộ lọc để mã hóa. Thành phần bộ lọc vòng kín 225 áp dụng các bộ lọc này dựa trên dữ liệu điều khiển bộ lọc. Các bộ lọc có

thể bao gồm bộ lọc giải khối, bộ lọc chặn nhiễu, bộ lọc SAO, và bộ lọc vòng lặp thích ứng. Các bộ lọc này có thể được áp dụng trong miền không gian/miền điểm ảnh (ví dụ, trên khối điểm ảnh được tái cấu trúc) hoặc trong miền tần số, tùy thuộc vào ví dụ.

Khi hoạt động như bộ mã hóa, khối hình ảnh tái cấu trúc được lọc, khối dư, và/hoặc khối dự đoán được chứa trong thành phần đệm hình ảnh giải mã 223 để sử dụng về sau trong ước tính chuyển động như đề cập đến bên trên. Khi hoạt động như bộ giải mã, thành phần đệm hình ảnh giải mã 223 lưu trữ và chuyển tiếp các khối được tái cấu trúc và lọc về phía màn hiển thị như một phần của đầu ra tín hiệu video. Thành phần đệm hình ảnh giải mã 223 có thể là thiết bị nhớ bất kỳ có khả năng lưu trữ các khối dự đoán, các khối dư, và/hoặc các khối hình ảnh được tái cấu trúc.

Thành phần định dạng tiêu đề và CABAC 231 nhận dữ liệu từ các thành phần khác nhau của hệ thống mã hóa 200 và mã hóa dữ liệu này thành dòng bit mã hóa để truyền về phía bộ giải mã. Cụ thể là, thành phần định dạng tiêu đề và CABAC 231 sinh ra các tiêu đề khác nhau để mã hóa dữ liệu điều khiển, như dữ liệu điều khiển chung và dữ liệu điều khiển bộ lọc. Ngoài ra, dữ liệu dự đoán, bao gồm dự đoán nội suy và dữ liệu chuyển động, cũng như dữ liệu dư dưới dạng của hệ số dữ liệu chuyển đổi đã được lượng tử hóa tất cả được mã hóa trong dòng bit. Dòng bit cuối cùng bao gồm tất cả thông tin được kỳ vọng bởi bộ giải mã để tái cấu trúc tín hiệu video ban đầu đã phân chia 201. Thông tin này có thể còn bao gồm chế độ dự đoán nội suy chỉ số bảngs (còn được gọi là các bảng ánh xạ từ mã), các định nghĩa của các ngữ cảnh mã hóa cho các khối khác nhau, các chỉ số của hầu hết các chế độ dự đoán nội suy có thể có, chỉ báo của thông tin phân chia, v.v. Dữ liệu này có thể được mã hóa bằng cách sử dụng mã entropy. Ví dụ, thông tin có thể được mã hóa bằng cách sử dụng mã chiều dài biến thiên thích ứng ngữ cảnh (CAVLC), CABAC, mã số nhị phân thích ứng ngữ cảnh dựa trên cú pháp (SBAC), mã entropy khả năng phân chia bên trong (PIPE), hoặc kỹ thuật mã entropy khác. Theo sau mã entropy, dòng bit được tạo mã có thể được truyền đến thiết bị khác (ví dụ, bộ giải mã video) hoặc đạt được cho hoạt động truyền hoặc truy tìm về sau.

FIG.3 là sơ đồ khối thể hiện bộ mã hóa video làm ví dụ 300. Bộ mã hóa video 300 có thể được sử dụng để thực hiện các chức năng mã hóa của hệ thống mã hóa 200 và/hoặc thực hiện các bước 103, 105, 107, và/hoặc 109 của phương pháp hoạt động

100. Bộ mã hóa 300 chia tín hiệu video đầu vào, tạo ra tín hiệu video đã phân chia 301, mà hầu như tương tự với tín hiệu video đã phân chia 201. Tín hiệu video đã phân chia 301 sau đó được nén và được mã hóa thành dòng bit bởi các thành phần của bộ mã hóa 300.

Cụ thể là, tín hiệu video đã phân chia 301 được chuyển tiếp đến thành phần dự đoán nội suy ảnh 317 để dự đoán nội suy. Thành phần dự đoán nội suy ảnh 317 có thể gần như tương tự với thành phần ước tính bên trong ảnh 215 và thành phần dự đoán nội suy ảnh 217. Tín hiệu video đã phân chia 301 cũng được chuyển tiếp đến thành phần bù chuyển động 321 để dự đoán ngoại suy dựa trên các khối tham chiếu trong thành phần đệm hình ảnh giải mã 323. Thành phần bù chuyển động 321 có thể hầu như tương tự với thành phần ước tính chuyển động 221 và thành phần bù chuyển động 219. Các khối dự đoán và các khối dư từ thành phần dự đoán nội suy ảnh 317 và thành phần bù chuyển động 321 được chuyển tiếp đến thành phần chuyển đổi và lượng tử hóa 313 để chuyển đổi và lượng tử hóa các khối dư. Thành phần chuyển đổi và lượng tử hóa 313 có thể hầu như tương tự với thành phần chuyển đổi tỷ lệ và lượng tử hóa 213. Các khối dư đã được chuyển đổi và lượng tử hóa và các khối dự đoán tương ứng (cùng với dữ liệu điều khiển kết hợp) được chuyển tiếp đến thành phần mã entropy 331 để mã hóa thành dòng bit. Thành phần mã entropy 331 có thể hầu như tương tự với thành phần định dạng tiêu đề và CABAC 231.

Các khối dư đã được chuyển đổi và lượng tử hóa và/hoặc các khối dự đoán tương ứng cũng được chuyển tiếp từ thành phần chuyển đổi và lượng tử hóa 313 đến thành phần chuyển đổi ngược và lượng tử hóa 329 để tái cấu trúc thành các khối tham chiếu để sử dụng bởi thành phần bù chuyển động 321. Thành phần chuyển đổi ngược và lượng tử hóa 329 có thể hầu như tương tự với thành phần chuyển đổi tỷ lệ và nghịch đảo 229. Các bộ lọc vòng kín trong thành phần bộ lọc vòng kín 325 cũng được áp dụng cho các khối dư và/hoặc các khối tái cấu trúc tham chiếu, tùy thuộc vào ví dụ. Thành phần bộ lọc vòng kín 325 có thể hầu như tương tự với thành phần phân tích điều khiển bộ lọc 227 và thành phần bộ lọc vòng kín 225. Thành phần bộ lọc vòng kín 325 có thể bao gồm nhiều bộ lọc như được mô tả đối với thành phần bộ lọc vòng kín 225. Các khối đã được lọc sau đó được lưu trữ trong thành phần đệm hình ảnh giải mã 323 để sử dụng làm các khối tham chiếu bởi thành phần bù chuyển động 321. Thành

phần đệm hình ảnh giải mã 323 có thể hầu như tương tự với thành phần đệm hình ảnh giải mã 223.

FIG.4 là sơ đồ khối thể hiện bộ giải mã video làm ví dụ 400. Bộ giải mã video 400 có thể được sử dụng để thực hiện các chức năng giải mã của hệ thống mã hóa 200 và/hoặc thực hiện các bước 111, 113, 115, và/hoặc 117 của phương pháp hoạt động 100. Bộ giải mã 400 nhận dòng bit, ví dụ từ bộ mã hóa 300, và sinh ra tín hiệu video đầu ra được tái cấu trúc dựa trên dòng bit để hiển thị cho người dùng đầu cuối.

Dòng bit được nhận bởi thành phần giải mã entropy 433. Thành phần giải mã entropy 433 được cấu hình để thực hiện sơ đồ giải mã entropy, như CAVLC, CABAC, SBAC, mã PIPE, hoặc các công nghệ mã entropy khác. Ví dụ, thành phần giải mã entropy 433 có thể sử dụng thông tin tiêu đề để cung cấp ngữ cảnh để biên dịch dữ liệu khác được mã hóa dưới dạng các từ mã trong dòng bit. Thông tin giải mã bao gồm bất kỳ thông tin kỳ vọng để giải mã tín hiệu video, như dữ liệu điều khiển chung, dữ liệu điều khiển bộ lọc, thông tin phân chia, dữ liệu chuyển động, dữ liệu dự đoán, và các hệ số chuyển đổi lượng tử hóa từ các khối dư. Các hệ số chuyển đổi được lượng tử hóa được chuyển tiếp đến thành phần chuyển đổi ngược và lượng tử hóa 429 để tái cấu trúc thành các khối dư. Thành phần chuyển đổi ngược và lượng tử hóa 429 có thể là tương tự với thành phần chuyển đổi ngược và lượng tử hóa 329.

Các khối dư tái cấu trúc và/hoặc các khối dự đoán được chuyển tiếp đến thành phần dự đoán nội suy ảnh 417 để tái cấu trúc thành các khối hình ảnh dựa trên các hoạt động dự đoán nội suy. Thành phần dự đoán nội suy ảnh 417 có thể là tương tự với thành phần ước tính bên trong ảnh 215 và thành phần dự đoán nội suy ảnh 217. Cụ thể là, thành phần dự đoán nội suy ảnh 417 sử dụng các chế độ dự đoán để định vị khối tham chiếu trên khung và áp dụng khối dư cho kết quả để tái cấu trúc các khối hình ảnh dự đoán nội suy. Các khối hình ảnh dự đoán nội suy tái cấu trúc và/hoặc các khối dư và dự đoán ngoại suy dữ liệu tương ứng được chuyển tiếp đến thành phần đệm hình ảnh giải mã 423 thông qua thành phần bộ lọc vòng kín 425, mà có thể hầu như tương tự với thành phần đệm hình ảnh giải mã 223 và thành phần bộ lọc vòng kín 225, một cách tương ứng). Thành phần bộ lọc vòng kín 425 lọc các khối hình ảnh được tái cấu trúc, các khối dư và/hoặc các khối dự đoán, và thông tin này được chứa trong thành phần đệm hình ảnh giải mã 423. Các khối hình ảnh được tái cấu trúc từ thành phần

đệm hình ảnh giải mã 423 được chuyển tiếp đến thành phần bù chuyển động 421 cho dự đoán ngoại suy. Thành phần bù chuyển động 421 có thể hầu như tương tự với thành phần ước tính chuyển động 221 và/hoặc thành phần bù chuyển động 219. Cụ thể là, thành phần bù chuyển động 421 sử dụng các vector chuyển động từ khối tham chiếu để tạo ra khối dự đoán và áp dụng khối dư cho kết quả để tái cấu trúc khối hình ảnh. Tạo thành các khối tái cấu trúc cũng có thể được chuyển tiếp thông qua thành phần bộ lọc vòng kín 425 đến thành phần đệm hình ảnh giải mã 423. Thành phần đệm hình ảnh giải mã 423 tiếp tục lưu trữ các khối hình ảnh được tái cấu trúc khác, mà có thể được tái cấu trúc thành các khung thông qua thông tin phân chia. Các khung này cũng có thể được đặt vào chuỗi. Chuỗi được gửi ra đến màn hiển thị làm tín hiệu video đầu ra được tái cấu trúc.

Theo nội dung bên trên, các công nghệ nén video thực hiện dự đoán không gian (trong ảnh) và/hoặc dự đoán thời gian (liên ảnh) để làm giảm hoặc loại bỏ sự dư thừa có hữu trong các chuỗi video. Đối với hoạt động tạo mã video dựa trên khối, lát cắt video (tức là, ảnh video hoặc phần của ảnh video) có thể được phân chia thành các khối video, mà cũng có thể được gọi là các khối cây, các khối cây mã hóa (CTB: coding tree block), các cụm cây mã hóa (CTU: coding tree unit), các cụm tạo mã (CU: coding unit) và/hoặc các nút tạo mã. Các khối video trong lát cắt được tạo mã trong ảnh (I) của ảnh được mã hóa nhờ sử dụng dự đoán không gian đối với các mẫu tham chiếu trong các khối lân cận trong ảnh giống nhau. Các khối video trong lát cắt được tạo mã liên ảnh (P hoặc B) của ảnh có thể sử dụng dự đoán không gian đối với các mẫu tham chiếu trong các khối lân cận trong ảnh giống nhau hoặc dự đoán thời gian đối với các mẫu tham chiếu trong các ảnh tham chiếu khác. Các ảnh có thể được gọi là các khung, và các ảnh tham chiếu có thể được gọi là các khung tham chiếu.

Dự đoán không gian hoặc thời gian tạo ra khối dự đoán cho khối cần được tạo mã. Dữ liệu dư thể hiện các khác nhau điểm ảnh giữa khối gốc cần được tạo mã và khối dự đoán. Khối được tạo mã liên ảnh được mã hóa theo vector chuyển động mà trở tới khối của các mẫu tham chiếu đang tạo ra khối dự đoán, và dữ liệu dư biểu thị sự chênh lệch giữa khối được tạo mã và khối dự đoán. Khối được tạo mã trong ảnh được mã hóa theo chế độ mã hóa trong ảnh và dữ liệu dư. Để nén thêm nữa, dữ liệu dư có thể được biến đổi từ miền điểm ảnh tới miền chuyển đổi, tạo ra các hệ số biến đổi dư,

mà sau đó có thể được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa, được bố trí ban đầu theo mảng hai chiều, có thể được quét để tạo ra vectơ một chiều của các hệ số biến đổi, và hoạt động tạo mã entropy có thể được áp dụng vào đạt được độ nén cao hơn nữa.

Hoạt động nén hình ảnh và video đã trải qua sự phát triển nhanh, dẫn tới các tiêu chuẩn mã hóa đa dạng. Các tiêu chuẩn tạo mã video này bao gồm ITU-T H.261, tổ chức tiêu chuẩn hóa quốc tế / Ủy ban kỹ thuật điện tử quốc tế (ISO/IEC) MPEG-1 phần 2, ITU-T H.262 hoặc ISO/IEC MPEG-2 phần 2, ITU-T H.263, ISO/IEC MPEG-4 phần 2, tạo mã video nâng cao (AVC), còn được biết đến là ITU-T H.264 hoặc ISO/IEC MPEG-4 phần 10, và tạo mã video hiệu suất cao (HEVC: High Efficiency Video Coding), còn được biết đến là ITU-T H.265 hoặc MPEG-H phần 2. AVC bao gồm các tiêu chuẩn mở rộng như tạo mã video thay đổi được (SVC), tạo mã video đa hình (MVC) và tạo mã video đa hình cộng với chiều sâu (MVC+D), và 3D AVC (3D-AVC). HEVC bao gồm các mở rộng như HEVC thay đổi được (SHVC), HEVC đa hình (MV-HEVC), và 3D HEVC (3D-HEVC).

Cũng có tiêu chuẩn mã hóa video mới, được gọi là tạo mã video vạn năng (VVC: Versatile Video Coding), được phát triển bởi nhóm chuyên gia video liên kết (JVET: joint video experts team) của ITU-T và ISO/IEC. Trong khi tiêu chuẩn VVC có một số dự thảo đang được thực hiện, một dự thảo đang được thực hiện (WD) của VVC cụ thể, cụ thể là B. Bross, J. Chen, và S. Liu, “Tạo mã video vạn năng (dự thảo 5),” JVET-N1001-v3, Hội nghị JVET lần thứ 13, 27/3/2019 (Dự thảo VVC 5) được viện dẫn ở đây.

Tạo mã video có lớp cũng được gọi là tạo mã video thay đổi được hoặc tạo mã video với khả năng thay đổi. Khả năng thay đổi trong tạo mã video thường được hỗ trợ bằng cách sử dụng các công nghệ tạo mã đa lớp. Dòng bit đa lớp bao gồm lớp gốc (BL: base layer) và một hoặc nhiều lớp tăng cường (EL: enhancement layer). Một ví dụ về các khả năng thay đổi bao gồm khả năng thay đổi không gian, khả năng thay đổi chất lượng/tín hiệu-với-ồn nhiễu (SNR: signal-to-noise), khả năng thay đổi đa màn hình, v.v. Khi công nghệ tạo mã đa lớp được sử dụng, ảnh hoặc một phần của chúng có thể được tạo mã (1) mà không cần sử dụng ảnh tham chiếu, tức là, nhờ sử dụng dự đoán trong ảnh; (2) bằng cách tham chiếu đến các ảnh tham chiếu mà nằm trong cùng

một lớp, tức là, nhờ sử dụng dự đoán liên ảnh; hoặc (3) bằng cách tham chiếu đến các ảnh tham chiếu mà nằm trong (các) lớp khác, tức là, sử dụng dự đoán liên lớp. Ảnh tham chiếu dùng cho dự đoán liên lớp của ảnh hiện tại được gọi là ảnh tham chiếu liên lớp (ILRP).

FIG.5 là hình vẽ dạng sơ đồ thể hiện ví dụ về dự đoán dựa trên lớp 500, ví dụ như được thực hiện để xác định các vector chuyển động (MVs) ở bước nén khối 105, bước giải mã khối 113, thành phần ước tính chuyển động 221, thành phần bù chuyển động 219, thành phần bù chuyển động 321, và/hoặc thành phần bù chuyển động 421. Dự đoán dựa trên lớp 500 là tương thích với dự đoán liên ảnh một chiều và/hoặc dự đoán liên ảnh hai chiều, nhưng cũng được thực hiện giữa các ảnh trong các lớp khác nhau.

Dự đoán dựa trên lớp 500 được phủ giữa các ảnh 511, 512, 513, và 514 và các ảnh 515, 516, 517, và 518 trong các lớp khác nhau. Theo ví dụ được thể hiện, các ảnh 511, 512, 513, và 514 là một phần của lớp N+1 532 và các ảnh 515, 516, 517, và 518 là một phần của lớp N 531. Lớp, như lớp N 531 và/hoặc lớp N+1 532, là nhóm của các ảnh mà tất cả được liên kết với giá trị đặc tính tương tự, như kích cỡ tương tự, chất lượng, độ phân giải, tỷ lệ tín hiệu với ồn nhiễu, khả năng, v.v. Theo ví dụ được thể hiện, lớp N+1 532 được liên kết với kích cỡ ảnh lớn hơn lớp N 531. Theo đó, các ảnh 511, 512, 513, và 514 trong lớp N+1 532 có kích cỡ ảnh lớn hơn (ví dụ, chiều cao và chiều rộng lớn hơn và do đó nhiều mẫu hơn) so với các ảnh 515, 516, 517, và 518 trong lớp N 531 trong ví dụ này. Tuy nhiên, các ảnh này có thể được tách biệt giữa lớp N+1 532 và lớp N 531 bởi các đặc tính khác. Trong khi chỉ hai lớp, lớp N+1 532 và lớp N 531, được thể hiện, tập hợp của các ảnh có thể được tách biệt thành số lượng bất kỳ của các lớp dựa trên các đặc tính được liên kết. Lớp N+1 532 và lớp N 531 cũng có thể được biểu thị bởi bộ nhận dạng lớp (ID). ID lớp là một mục của dữ liệu mà được liên kết với ảnh và biểu thị ảnh là một phần của lớp được chỉ báo. Theo đó, mỗi ảnh 511-518 có thể được liên kết với lớp tương ứng ID để chỉ báo lớp nào trong số lớp N+1 532 hoặc lớp N 531 bao gồm ảnh tương ứng.

Các ảnh 511-518 trong các lớp khác nhau 531-532 được tạo cấu hình để được hiển thị theo cách thay thế. Như vậy, các ảnh 511-518 trong các lớp khác nhau 531-532 có thể chia sẻ bộ nhận dạng thời gian giống nhau (ID) và có thể được chứa trong

AU giống nhau. Như được sử dụng ở đây, AU là tập hợp của một hoặc nhiều ảnh được tạo mã được kết hợp với cùng thời gian hiển thị cho việc xuất ra từ DPB. Ví dụ, bộ giải mã có thể giải mã và hiển thị ảnh 515 ở thời gian hiển thị hiện tại nếu ảnh nhỏ hơn được yêu cầu hoặc bộ giải mã có thể giải mã và hiển thị ảnh 511 ở thời gian hiển thị hiện tại nếu ảnh lớn hơn được yêu cầu. Như vậy, các ảnh 511-514 ở N+1 532 lớp cao hơn chứa dữ liệu hình ảnh hầu như giống với các ảnh tương ứng 515-518 ở lớp dưới N 531 (bất kể độ chênh lệch trong kích cỡ ảnh). Cụ thể là, ảnh 511 chứa dữ liệu hình ảnh hầu như giống với ảnh 515, ảnh 512 chứa dữ liệu hình ảnh hầu như giống với ảnh 516, v.v.

Các ảnh 511-518 có thể được tạo mã nhờ tham chiếu đến các ảnh khác 511-518 trong lớp giống nhau N 531 hoặc N+1 532. Việc tạo mã ảnh khi tham chiếu đến ảnh khác trong lớp giống nhau dẫn tới dự đoán liên ảnh 523, vốn là dự đoán liên ảnh một chiều và/hoặc dự đoán liên ảnh hai chiều tương thích. Dự đoán liên ảnh 523 được minh họa bởi các mũi tên nét liền. Ví dụ, ảnh 513 có thể được tạo mã bằng cách sử dụng dự đoán liên ảnh 523 nhờ sử dụng một hoặc hai trong số các ảnh 511, 512, và/hoặc 514 trong lớp N+1 532 làm tham chiếu, trong đó một ảnh được tham chiếu cho dự đoán liên ảnh một chiều và/hoặc hai ảnh được tham chiếu cho dự đoán liên ảnh hai chiều. Ngoài ra, ảnh 517 có thể được tạo mã bằng cách sử dụng dự đoán liên ảnh 523 nhờ sử dụng một hoặc hai trong số các ảnh 515, 516, và/hoặc 518 trong lớp N 531 làm tham chiếu, trong đó một ảnh được tham chiếu cho dự đoán liên ảnh một chiều và/hoặc hai ảnh được tham chiếu cho dự đoán liên ảnh hai chiều. Khi ảnh được sử dụng như một tham chiếu cho ảnh khác trong lớp giống nhau khi thực hiện dự đoán liên ảnh 523, ảnh có thể được gọi là ảnh tham chiếu. Ví dụ, ảnh 512 có thể là ảnh tham chiếu được sử dụng để tạo mã ảnh 513 theo dự đoán liên ảnh 523. Dự đoán liên ảnh 523 cũng có thể được gọi là dự đoán bên trong lớp trong ngữ cảnh đa lớp. Như vậy, dự đoán liên ảnh 523 là cơ chế tạo mã các mẫu của ảnh hiện tại nhờ tham chiếu đến các mẫu được chỉ báo trong ảnh tham chiếu vốn khác với ảnh hiện tại trong đó ảnh tham chiếu và ảnh hiện tại nằm trong cùng một lớp.

Các ảnh 511-518 cũng có thể được tạo mã nhờ tham chiếu đến các ảnh khác 511-518 trong các lớp khác nhau. Quá trình này được biết là dự đoán liên lớp 521, và được minh họa bởi các mũi tên bằng nét đứt. Dự đoán liên lớp 521 là cơ chế tạo mã

các mẫu của ảnh hiện tại nhờ tham chiếu đến các mẫu được chỉ báo trong ảnh tham chiếu trong đó ảnh hiện tại và ảnh tham chiếu nằm trong các lớp khác nhau và do đó có các ID lớp khác nhau. Ví dụ, ảnh ở lớp dưới N 531 có thể được sử dụng làm ảnh tham chiếu để tạo mã ảnh tương ứng ở lớp cao hơn N+1 532. Theo cụ thể ví dụ, ảnh 511 có thể được tạo mã nhờ tham chiếu đến ảnh 515 theo dự đoán liên lớp 521. Trong trường hợp như vậy, ảnh 515 được sử dụng như ảnh tham chiếu liên lớp. Ảnh tham chiếu liên lớp là ảnh tham chiếu dùng cho dự đoán liên lớp 521. Trong hầu hết các trường hợp, dự đoán liên lớp 521 được hạn chế sao cho ảnh hiện tại, như ảnh 511, có thể chỉ sử dụng (các) ảnh tham chiếu liên lớp mà được chứa trong AU giống nhau và nằm ở lớp dưới, như ảnh 515. Khi nhiều lớp (ví dụ, nhiều hơn hai) là khả dụng, dự đoán liên lớp 521 có thể mã hóa/giải mã ảnh hiện tại dựa trên (nhiều) ảnh tham chiếu liên lớp ở các mức thấp hơn ảnh hiện tại.

Bộ mã hóa video có thể sử dụng dự đoán dựa trên lớp 500 để mã hóa các ảnh 511-518 thông qua nhiều kết hợp khác nhau và/hoặc các hoán vị của dự đoán liên ảnh 523 và dự đoán liên lớp 521. Ví dụ, ảnh 515 có thể được tạo mã theo dự đoán trong ảnh. Sau đó các ảnh 516-518 có thể được tạo mã theo dự đoán liên ảnh 523 bằng cách sử dụng ảnh 515 làm ảnh tham chiếu. Ngoài ra, ảnh 511 có thể được tạo mã theo dự đoán liên lớp 521 bằng cách sử dụng ảnh 515 làm ảnh tham chiếu liên lớp. Sau đó các ảnh 512-514 có thể được tạo mã theo dự đoán liên ảnh 523 bằng cách sử dụng ảnh 511 làm ảnh tham chiếu. Như vậy, ảnh tham chiếu có thể có cả vai trò như ảnh tham chiếu một lớp và ảnh tham chiếu liên lớp cho các cơ chế tạo mã khác nhau. Bằng cách tạo mã các ảnh N+1 532 lớp cao hơn dựa trên các ảnh N 531 lớp dưới, N+1 532 lớp cao hơn có thể tránh được việc sử dụng dự đoán trong ảnh, vốn có hiệu suất tạo mã thấp hơn nhiều so với dự đoán liên ảnh 523 và dự đoán liên lớp 521. Như vậy, hiệu suất tạo mã thấp của dự đoán trong ảnh có thể bị giới hạn ở các ảnh có chất lượng nhỏ nhất/thấp nhất, và do đó bị giới hạn ở việc tạo mã lượng dữ liệu video nhỏ nhất. Các ảnh được sử dụng làm các ảnh tham chiếu và/hoặc các ảnh tham chiếu liên lớp có thể được chỉ báo trong các mục nhập của (các) danh sách ảnh tham chiếu chứa trong cấu trúc danh sách ảnh tham chiếu.

Mỗi AU 506 trên Fig.5 có thể chứa một số ảnh. Ví dụ, một AU 506 có thể chứa các ảnh 511 và 515. AU 506 khác có thể chứa các ảnh 512 và 516. Thực vậy, mỗi AU

506 là tập hợp của một hoặc nhiều ảnh được tạo mã được kết hợp với cùng thời gian hiển thị (ví dụ, ID thời gian giống nhau) cho việc xuất ra từ bộ đệm hình ảnh đã được giải mã (DPB: decoded picture buffer) (ví dụ, để hiển thị cho người dùng). Mỗi bộ tách cụm truy nhập (AUD) 508 là bộ chỉ báo hoặc cấu trúc dữ liệu được sử dụng để chỉ báo sự bắt đầu của AU (ví dụ AU 506) hoặc đường bao giữa các AU.

Theo một phương án, một AU (ví dụ, AU 506) chứa ảnh điểm truy nhập ngẫu nhiên trong ảnh (IRAP: intra random access point) (ví dụ, ảnh 515) trong một lớp (ví dụ, lớp N 531) và ảnh không có IRAP (ví dụ, ảnh 511) trong lớp khác (ví dụ, lớp N+1 532). Tức là, AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau. Trong khi AU được thể hiện chứa một ảnh IRAP và một ảnh không có IRAP, số lượng các ảnh IRAP và không có IRAP khác nhau có thể có mặt trong AU trong các ứng dụng thực tế. Các ảnh IRAP tiếp tục được mô tả chi tiết dưới đây.

Các họ tạo mã video H.26x trước đó đã cung cấp sự hỗ trợ cho khả năng thay đổi trong các tiêu sử tách biệt với các tiêu sử cho việc tạo mã một lớp đơn. Hoạt động tạo mã video thay đổi được (SVC) là sự mở rộng thay đổi được của AVC/H.264 mà tạo ra hỗ trợ cho các khả năng thay đổi không gian, thời gian và chất lượng. Đối với SVC, cờ được báo hiệu trong mỗi khối macro (MB: macroblock) trong các ảnh EL để chỉ báo xem EL MB được dự đoán nhờ sử dụng khối được sắp xếp tạm thời từ lớp dưới. Việc dự đoán từ khối được sắp xếp tạm thời có thể bao gồm kết cấu, các vector chuyển động, và/hoặc các chế độ tạo mã. Các cách thực hiện của SVC không thể sử dụng lại một cách trực tiếp các cách thực hiện H.264/AVC không hiệu chỉnh trong thiết kế của chúng. Cú pháp khối macro SVC EL và quy trình giải mã khác với cú pháp H.264/AVC và quy trình giải mã.

HEVC thay đổi được (SHVC) là sự mở rộng của tiêu chuẩn HEVC/H.265 mà cung cấp sự hỗ trợ cho các khả năng thay đổi không gian và chất lượng, HEVC đa hình (MV-HEVC) là sự mở rộng của HEVC/H.265 mà cung cấp sự hỗ trợ cho khả năng thay đổi đa hình, và 3D HEVC (3D-HEVC) là sự mở rộng của HEVC/H.264 mà cung cấp sự hỗ trợ cho việc tạo mã video ba chiều (3D) vốn cải tiến hơn và hiệu quả hơn so với MV-HEVC. Chú ý rằng khả năng thay đổi thời gian được bao hàm như một phần gắn liền của phương tiện mã hóa HEVC một lớp đơn. Thiết kế của đa lớp sự mở rộng của HEVC sử dụng ý tưởng mà trong đó các ảnh được giải mã dùng cho dự

đoán liên lớp chỉ đến từ cụm truy nhập (AU) giống nhau và được xử lý như các ảnh tham chiếu dài hạn (LTRP: long-term reference picture), và các chỉ số tham chiếu được ấn định trong (các) danh sách ảnh tham chiếu cùng với các ảnh tham chiếu theo thời gian khác trong lớp hiện tại. Dự đoán liên lớp (ILP: Inter-layer prediction) đạt được ở mức cụm dự đoán (PU) bằng cách thiết lập giá trị của chỉ số tham chiếu để tham chiếu đến (các) ảnh tham chiếu liên lớp trong (các) danh sách ảnh tham chiếu.

Đáng chú ý là, cả việc lấy mẫu ảnh tham chiếu và việc gọi các đặc trưng khả năng thay đổi không gian cho việc lấy mẫu của ảnh tham chiếu hoặc một phần của chúng. Việc lấy mẫu ảnh tham chiếu (RPR) có thể được thực hiện ở mức ảnh hoặc mức khối tạo mã. Tuy nhiên, khi RPR được gọi là đặc trưng tạo mã, nó là đặc trưng cho việc tạo mã một lớp đơn. Thậm chí như vậy, có thể hoặc thậm chí có thể ưu tiên từ quan điểm thiết kế phương tiện mã hóa để sử dụng bộ lọc mẫu giống nhau cho cả đặc trưng RPR của việc tạo mã một lớp đơn và đặc trưng khả năng thay đổi không gian cho việc tạo mã đa lớp.

Bản thảo VVC cuối cùng hỗ trợ việc tạo mã video có lớp. Dòng bit VVC có thể bao gồm nhiều lớp. Tất cả các lớp có thể là độc lập với nhau, tức là, mỗi lớp được tạo mã mà không sử dụng dự đoán liên lớp (ILP). Trong trường hợp này, các lớp còn được gọi là các lớp đồng truyền. Cũng có thể có khả năng là một số lớp trong số các lớp được tạo mã nhờ sử dụng ILP. Còn trong VPS được sử dụng để chỉ báo xem các lớp là các lớp đồng truyền hoặc xem một số lớp có sử dụng ILP hay không. Khi một số lớp sử dụng ILP, quan hệ phụ thuộc lớp trong số các lớp cũng được báo hiệu trong VPS.

Không giống như SHVC và MV-HEVC, bản thảo VVC cuối cùng không chỉ định các OLS. OLS là tập hợp của các lớp mà một hoặc nhiều lớp được định rõ cho chúng là các lớp đầu ra. Lớp đầu ra là lớp của tập hợp lớp đầu ra mà được xuất ra.

Trong bản thảo VVC cuối cùng, khi các lớp là các lớp đồng truyền, thì chỉ định rõ ràng chỉ một lớp có thể được chọn để giải mã và xuất ra. Trong bản thảo VVC cuối cùng, khi một số lớp sử dụng ILP, tất cả các lớp trong dòng bit được định rõ sẽ được giải mã nhưng chỉ các lớp nhất định được định rõ sẽ là các lớp đầu ra. Các lớp đầu ra có thể được chỉ báo để 1) chỉ là lớp cao nhất, 2) là tất cả các lớp, hoặc 3) là lớp cao nhất cộng tập hợp của các lớp dưới được chỉ báo.

FIG.6 là sự biểu diễn 600 của quan hệ giữa ảnh điểm truy nhập ngẫu nhiên trong ảnh (IRAP: intra random access point) 602 tương đối với các ảnh phía trước 604 và các ảnh phía sau 606 trong thứ tự giải mã 608 và thứ tự biểu diễn 610 (tức là, thứ tự xuất). Theo một phương án, ảnh IRAP 602 được gọi là ảnh truy nhập ngẫu nhiên sạch (CRA: clean random access) hoặc là ảnh làm mới bộ giải mã tức thời (IDR: instantaneous decoder refresh) với ảnh có thể giải mã truy nhập ngẫu nhiên (RADL). Trong HEVC, các ảnh IDR, các ảnh CRA, và các ảnh truy nhập liên kết gãy (BLA: Broken Link Access) tất cả được xem là các ảnh IRAP 602. Đối với VVC, trong hội thảo JVET lần 12 vào tháng 10, 2018, VVC được đồng ý có cả hai ảnh IDR và CRA làm các ảnh IRAP. Theo một phương án, các ảnh truy nhập liên kết gãy (BLA) và làm mới bộ giải mã dần dần (GDR) cũng có thể được xem là các ảnh IRAP. Quy trình giải mã cho chuỗi video đã được tạo mã luôn bắt đầu ở IRAP. Các ảnh IRAP cung cấp điểm truy nhập ngẫu nhiên (RAP).

Như được thể hiện trên FIG.6, các ảnh phía trước 604 (ví dụ, các ảnh 2 và 3) theo sau ảnh IRAP 602 trong thứ tự giải mã 608, nhưng ở trước ảnh IRAP 602 trong thứ tự biểu diễn 610. Ảnh phía sau 606 theo sau ảnh IRAP 602 trong cả thứ tự giải mã 608 và thứ tự biểu diễn 610. Trong khi hai ảnh phía trước 604 và một ảnh phía sau 606 được minh họa trên FIG.6, người có hiểu biết trung bình trong lĩnh vực tương ứng sẽ hiểu rõ rằng nhiều hoặc ít hơn các ảnh phía trước 604 và/hoặc các ảnh phía sau 606 có thể có mặt trong thứ tự giải mã 608 và thứ tự biểu diễn 610 trong các ứng dụng thực tế.

Các ảnh phía trước 604 trên FIG.6 được chia thành hai loại, cụ thể là hàng đầu được bỏ qua truy nhập ngẫu nhiên (RASL) và RADL. Khi bắt đầu việc giải mã với ảnh IRAP 602 (ví dụ, ảnh 1), thì ảnh RADL (ví dụ, ảnh 3) có thể được giải mã một cách thích hợp; tuy nhiên, ảnh RASL (ví dụ, ảnh 2) không thể được giải mã một cách thích hợp. Do vậy, ảnh RASL được loại bỏ. Nhờ sự phân biệt giữa các ảnh RADL và RASL, kiểu ảnh phía trước 604 được liên kết với ảnh IRAP 602 nên được xác định là RADL hoặc RASL cho việc tạo mã hiệu quả và thích hợp. Trong HEVC, khi các ảnh RASL và RADL được thể hiện, thì có hạn chế là đối với các ảnh RASL và RADL mà được liên kết với ảnh IRAP 602 giống nhau, các ảnh RASL sẽ ở phía trước các ảnh RADL theo thứ tự biểu diễn 610.

Ảnh IRAP 602 cung cấp hai chức năng/lợi ích quan trọng sau đây. Trước tiên, sự có mặt của ảnh IRAP 602 chỉ báo rằng quy trình giải mã có thể bắt đầu từ ảnh đó. Chức năng này cho phép đặc trưng truy nhập ngẫu nhiên trong đó quy trình giải mã bắt đầu ở vị trí đó trong dòng bit, không cần sự bắt đầu của dòng bit, miễn là ảnh IRAP 602 có mặt ở vị trí đó. Thứ hai, sự có mặt của ảnh IRAP 602 nạp lại quy trình giải mã sao cho ảnh được tạo mã bắt đầu ở ảnh IRAP 602, ngoại trừ các ảnh RASL, được tạo mã mà không cần sự tham chiếu bất kỳ tới các ảnh trước đó. Do đó việc có ảnh IRAP 602 thể hiện trong dòng bit sẽ dừng lỗi bất kỳ mà có thể xảy ra trong khi giải mã các ảnh được tạo mã trước khi ảnh IRAP 602 lan truyền tới ảnh IRAP 602 và các ảnh nào mà theo sau ảnh IRAP 602 theo thứ tự giải mã 608.

Trong khi các ảnh IRAP 602 cung cấp các chức năng quan trọng, chúng có thể gây bất lợi cho hiệu quả nén. Sự có mặt của ảnh IRAP 602 gây ra đột biến điện trong tốc độ bit. Sự bất lợi này đối với hiệu quả nén là do hai nguyên nhân. Một là, do ảnh IRAP 602 là ảnh được dự đoán trong ảnh, ảnh này tự nó sẽ yêu cầu tương đối nhiều bit hơn để biểu diễn khi so sánh với các ảnh khác (ví dụ, các ảnh phía trước 604, các ảnh phía sau 606) vốn là các ảnh được dự đoán liên ảnh. Hai là, do sự có mặt của ảnh IRAP 602 phá vỡ sự dự đoán thời gian (điều này là do bộ giải mã sẽ làm mới quy trình giải mã, trong đó một trong số hoạt động của quy trình giải mã cho sự giải mã này sẽ loại bỏ các ảnh tham chiếu trước đó trong bộ đệm hình ảnh đã được giải mã (DPB: decoded picture buffer), ảnh IRAP 602 khiến cho việc tạo mã các ảnh mà theo sau ảnh IRAP 602 theo thứ tự giải mã 608 kém hiệu quả hơn (tức là, cần nhiều bit hơn để biểu diễn) do chúng có ít các ảnh tham chiếu hơn cho việc tạo mã dự đoán liên ảnh của chúng.

Trong số các loại ảnh được xem là các ảnh IRAP 602, ảnh IDR trong HEVC có báo hiệu và sự suy ra khác nhau khi so với các loại ảnh khác. Một số khác nhau là như sau.

Đối với báo hiệu và sự suy ra của giá trị số đếm thứ tự ảnh (POC) của ảnh IDR, phần bit đáng kể nhất (MSB) của POC không được suy ra từ ảnh chính trước đó nhưng được thiết lập theo cách đơn giản cho bằng 0.

Đối với thông tin báo hiệu được yêu cầu cho việc quản lý ảnh tham chiếu, tiêu đề lát của ảnh IDR không chứa thông tin cần được báo hiệu để hỗ trợ sự quản lý ảnh

tham chiếu. Đối với các kiểu ảnh khác (tức là, CRA, Trailing, truy nhập lớp con thời gian (TSA), v.v.), thông tin như nhóm ảnh tham chiếu (RPS) được mô tả dưới đây hoặc các dạng khác của thông tin tương tự (ví dụ, các danh sách ảnh tham chiếu) được yêu cầu đối với các quy trình đánh dấu ảnh tham chiếu (tức là, quy trình để xác định trạng thái của các ảnh tham chiếu trong bộ đệm hình ảnh đã được giải mã (DPB: decoded picture buffer), hoặc được sử dụng cho tham chiếu và không được sử dụng cho tham chiếu). Tuy nhiên, đối với ảnh IDR, thông tin này không cần được báo hiệu do sự có mặt của chỉ báo IDR rằng quy trình giải mã sẽ đánh dấu một cách đơn giản tất cả các ảnh tham chiếu trong DPB là không được sử dụng để tham chiếu.

Trong HEVC và VVC, các ảnh các ảnh IRAP và dẫn được cho các loại cụm NAL (NUT: NAL unit type) khác nhau sao cho chúng có thể được nhận dạng một cách dễ dàng bởi các ứng dụng mức hệ thống. Ví dụ, bộ nội video cần hiểu các loại ảnh đã được tạo mã mà không phải hiểu quá nhiều chi tiết của thành phần cú pháp trong dòng bit được tạo mã, cụ thể là để xác định các ảnh IRAP từ các ảnh không có IRAP và để xác định các ảnh phía trước, bao gồm xác định các ảnh RASL và RADL, từ các ảnh phía sau. Các ảnh phía sau là các ảnh mà được liên kết với ảnh IRAP và theo sau ảnh IRAP theo thứ tự xuất. Ảnh được liên kết với ảnh IRAP cụ thể theo sau ảnh IRAP cụ thể theo thứ tự giải mã và ở trước ảnh IRAP bất kỳ khác theo thứ tự giải mã. Đối với hoạt động này, việc cho các ảnh IRAP và ảnh phía trước của chúng sở hữu kiểu cụm NAL có thể giúp được các ứng dụng này.

Trong HEVC, mỗi ảnh có giá trị POC được ấn định cho nó, được biểu thị là PicOrderCntVal. POC có ba cách sử dụng chính: để xác định duy nhất các ảnh, để chỉ báo vị trí đầu ra tương đối với các ảnh khác trong CVS giống nhau, và để thực hiện tỷ lệ vector chuyển động trong quy trình giải mã lớp tạo mã video mức thấp (VCL). Tất cả các ảnh trong CVS giống nhau có giá trị POC duy nhất. Các ảnh từ các CVS khác nhau có thể chia sẻ giá trị POC giống nhau, nhưng các ảnh vẫn có thể được xác định duy nhất do không có các khả năng để trộn các ảnh từ một CVS với ảnh bất kỳ của CVS khác. Các khe hở trong các giá trị POC được cho phép trong CVS. Tức là, sự chênh lệch giá trị POC giữa hai ảnh liên tiếp theo thứ tự xuất có thể chênh lệch nhiều hơn một (và lượng chênh lệch thực tế mà các giá trị POC cho các ảnh liên tiếp có thể chênh lệch theo đó có thể thay đổi tùy ý).

Trong HEVC, giá trị POC của ảnh được báo hiệu nhờ sử dụng từ mã trong tiêu đề lát. Phạm vi của các giá trị POC được phép nằm trong khoảng từ -2^{31} đến $2^{31}-1$, để tiết kiệm các bit trong tiêu đề lát, chỉ các bit ít đáng kể nhất của giá trị POC (POC LSB) được báo hiệu. Số lượng các bit để sử dụng cho POC LSB có thể nằm trong khoảng từ 4 đến 16, và được báo hiệu trong SPS. Phần còn lại của các bit là các bit có giá trị POC đáng kể nhất (POC MSB). Do chỉ POC LSB được báo hiệu trong tiêu đề lát, POC MSB cho ảnh hiện tại được suy ra từ ảnh trước đó, gọi là prevTid0Pic. Nhằm suy ra POC để làm việc theo cách giống nhau ngay cả khi các ảnh được loại bỏ, prevTid0Pic được chọn là ảnh trước đó gần nhất của lớp thời gian 0 mà không là ảnh RASL, ảnh RADL, hoặc ảnh không tham chiếu lớp con. Bộ giải mã suy ra giá trị POC MSB bằng cách so sánh giá trị POC của ảnh hiện tại với giá trị POC của ảnh prevTid0Pic.

Không may là, các thiết kế tạo mã hiện có có các nhược điểm. Ví dụ, trong SHVC và MV-HEVC, các ảnh IRAP được trộn và không có IRAP có thể được bao hàm bên trong cùng cụm truy nhập (ví dụ, ảnh 511 và 515 trong AU 506 trên FIG.5), nhưng thiết kế là rất phức tạp. Thực vậy, quy trình giải mã chung cho thiết kế này được chỉ định rõ trong SHVC và MV-HEVC nhờ sử dụng một số trang của văn bản trong mệnh đề F.8.1 của tiêu chuẩn HEVC cuối cùng. Văn bản bao gồm việc sử dụng nhiều cờ/biến số/danh sách toàn cầu, bao gồm NoClasOutputFlag, LayerInitializedFlag, LayerResetFlag, FirstPicInLayerDecodedFlag, NumPredictedLayers, IdPredictedLayer, IdDirectRefLayer, NumDirectRefLayers, NumActiveRefLayerPics, NumActiveRefLayerPics0, NumActiveRefLayerPics1, v.v. Nhiều vị trí khác trong SHVC và MV-HEVC mà tạo ra hỗ trợ cho các ảnh IRAP và không có IRAP được trộn bên trong cụm truy nhập cũng là phức tạp. Sự phức tạp này bao gồm quy trình suy ra POC.

Được bộc lộ ở đây là các công nghệ mà sự suy ra đơn giản hóa số đếm thứ tự ảnh (POC) trong tạo mã video đa lớp khi cụm truy nhập (AU) giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị bit đáng kể nhất (MSB) POC được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh

không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit (thay vì được suy ra như trong HEVC). Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

FIG.7 minh họa dòng bit video 750 được tạo cấu hình để thực hiện kỹ thuật làm mới việc giải mã (GDR) dần dần 700. Như được sử dụng ở đây, dòng bit video 750 cũng có thể được gọi là dòng bit video được tạo mã, dòng bit, hoặc các biến thể của chúng. Như được thể hiện trên FIG.7, dòng bit 750 bao gồm tập thông số chuỗi (SPS) 752, tập thông số ảnh (PPS) 754, tiêu đề lát 756, và dữ liệu hình ảnh 758.

SPS 752 chứa dữ liệu chung cho tất cả các ảnh trong chuỗi của các ảnh (SOP). Ngược lại, PPS 754 chứa dữ liệu chung cho toàn bộ ảnh. Tiêu đề lát 756 chứa thông tin về lát cắt hiện tại chẳng hạn như loại lát cắt, mà các ảnh tham chiếu của chúng sẽ được sử dụng, v.v.. SPS 752 và PPS 754 có thể được gọi chung là tập thông số. SPS 752, PPS 754, và tiêu đề lát 756 là các loại của các cụm lớp trừu tượng mạng (NAL: network abstraction layer). Cụm NAL là cấu trúc cú pháp chứa chỉ báo về loại của dữ liệu cần tuân theo (ví dụ, dữ liệu video được tạo mã). Các cụm NAL được phân loại thành lớp tạo mã video (VCL) và các cụm không phải VCL NAL. Các cụm VCL NAL chứa dữ liệu mà thể hiện các giá trị của các mẫu trong các ảnh video, và các cụm không phải VCL NAL chứa thông tin bổ sung kết hợp bất kỳ như các tập thông số (dữ liệu quan trọng mà có thể áp dụng cho số lượng các cụm VCL NAL) và thông tin tăng cường bổ sung (thông tin định thời và dữ liệu bổ sung khác mà có thể tăng cường khả năng sử dụng của tín hiệu video được giải mã nhưng là không cần thiết cho việc giải mã các giá trị của các mẫu trong các ảnh video). Người có hiểu biết trung bình trong

lĩnh vực tương ứng sẽ hiểu rõ rằng dòng bit 750 có thể chứa các thông số và thông tin khác trong các ứng dụng thực tế.

Dữ liệu hình ảnh 758 trên FIG.7 bao gồm dữ liệu được liên kết với các hình ảnh hoặc video được mã hóa hoặc được giải mã. Dữ liệu hình ảnh 758 có thể được gọi một cách đơn giản là phần tải dữ liệu hoặc dữ liệu được mang trong dòng bit 750. Theo một phương án, dữ liệu hình ảnh 758 bao gồm CVS 708 (hoặc CLVS) chứa ảnh GDR 702, một hoặc nhiều ảnh phía sau 704, và ảnh điểm hồi phục 706. Theo một phương án, ảnh GDR 702 được gọi là ảnh bắt đầu CVS (CVSS). CVS 708 là chuỗi video đã được tạo mã cho mỗi chuỗi video lớp được tạo mã (CLVS) trong dòng bit video 750. Đáng chú ý là, CVS và CLVS là giống nhau khi dòng bit video 750 bao gồm một lớp. CVS và CLVS chỉ khác nhau khi dòng bit video 750 bao gồm nhiều lớp. Theo một phương án, các ảnh phía sau 704 có thể được coi là một dạng của ảnh GDR do chúng ở trước ảnh điểm hồi phục 706 trong chu kỳ GDR.

Theo một phương án, ảnh GDR 702, các ảnh phía sau 704, và ảnh điểm hồi phục 706 có thể xác định chu kỳ GDR trong CVS 708. Theo một phương án, thứ tự giải mã bắt đầu với ảnh GDR 702, tiếp tục với các ảnh phía sau 704, và sau đó thực hiện đến ảnh hồi phục 706.

CVS 708 là loạt các ảnh (hoặc các phần của chúng) bắt đầu với ảnh GDR 702 và bao gồm tất cả các ảnh (hoặc các phần của chúng) lên tới, nhưng không bao gồm, ảnh GDR tiếp theo hoặc cho đến khi kết thúc dòng bit 750. Chu kỳ GDR là loạt các ảnh bắt đầu với ảnh GDR 702 và bao gồm tất cả các ảnh lên tới và bao gồm ảnh điểm hồi phục 706. Quy trình giải mã cho CVS 708 luôn bắt đầu ở ảnh GDR 702.

Như được thể hiện trên FIG.7, kỹ thuật GDR 700 hoặc nguyên lý làm việc trên loạt các ảnh bắt đầu với ảnh GDR 702 và kết thúc với ảnh điểm hồi phục 706. Ảnh GDR 702 chứa vùng được làm mới/sạch 710 chứa các khối mà tất cả được tạo mã nhờ sử dụng dự đoán trong ảnh (tức là, các khối được dự đoán trong ảnh) và vùng không được làm mới/bản 712 chứa các khối mà tất cả được tạo mã nhờ sử dụng dự đoán liên ảnh (tức là, các khối được dự đoán liên ảnh).

Ảnh phía sau 704 liền kề ngay với ảnh GDR 702 chứa vùng được làm mới/sạch 710 có phần thứ nhất 710A được tạo mã nhờ sử dụng dự đoán trong ảnh và phần thứ hai 710B được tạo mã nhờ sử dụng dự đoán liên ảnh. Phần thứ hai 710B

được tạo mã bằng cách tham chiếu vùng được làm mới/sạch 710 ví dụ bằng, ảnh trước đó bên trong chu kỳ GDR của CVS 708. Như được thể hiện, vùng được làm mới/sạch 710 của các ảnh phía sau 704 mở rộng khi quy trình tạo mã dịch chuyển hoặc phát triển theo hướng nhất quán (ví dụ, từ trái sang phải), mà theo cách tương ứng co lại vùng không được làm mới/bản 712. Cuối cùng, ảnh điểm hồi phục 706, mà chứa chỉ vùng được làm mới/sạch 710, được thu từ quy trình tạo mã. Đáng chú ý là, và như sẽ được tiếp tục mô tả bên dưới, phần thứ hai 710B của vùng được làm mới/sạch 710, mà được tạo mã như các khối được dự đoán liên ảnh, có thể chỉ viện dẫn đến vùng được làm mới/sạch 710 trong ảnh tham chiếu.

Như được thể hiện trên FIG.7, các lát cắt của ảnh GDR 702, các ảnh phía sau 704, và ảnh điểm hồi phục 706 trong CVS 708 mỗi ảnh được chứa bên trong cụm VCL NAL 730 của chính chúng.

Theo một phương án, cụm VCL NAL 730 chứa ảnh GDR 702 trong CVS 708 có kiểu cụm GDR NAL (GDR_NUT). Tức là, theo một phương án cụm VCL NAL 730 chứa ảnh GDR 702 trong CVS 708 có kiểu cụm NAL duy nhất của chính nó tương đối với các ảnh phía sau 704 và ảnh điểm hồi phục 706. Theo một phương án, GDR_NUT cho phép dòng bit 750 bắt đầu với ảnh GDR 702 thay vì dòng bit 750 có để bắt đầu với ảnh điểm truy nhập ngẫu nhiên trong ảnh (IRAP: intra random access point). Việc chỉ định cụm VCL NAL 730 của ảnh GDR 702 làm GDR_NUT có thể chỉ báo cho, ví dụ, bộ giải mã rằng cụm VCL NAL 730 ban đầu trong CVS 708 chứa ảnh GDR 702. Theo một phương án, ảnh GDR 702 là ảnh ban đầu trong CVS 708. Theo một phương án, ảnh GDR 702 là ảnh ban đầu trong chu kỳ GDR.

FIG.8 là hình vẽ dạng giản lược thể hiện việc tìm kiếm chuyển động không mong muốn 800 khi sử dụng sự hạn chế bộ mã hóa để hỗ trợ GDR. Như được thể hiện, chuyển động tìm kiếm 800 thể hiện ảnh hiện tại 802 và ảnh tham chiếu 804. Mỗi ảnh hiện tại 802 và ảnh tham chiếu 804 bao gồm vùng được làm mới 806 được tạo mã với dự đoán trong ảnh, vùng được làm mới 808 được tạo mã với dự đoán liên ảnh, và vùng không được làm mới 810. Vùng được làm mới 806, vùng được làm mới 808, và vùng không được làm mới 810 là tương tự với phần thứ nhất 710A của vùng được làm mới/sạch 710, phần thứ hai 710B của vùng được làm mới/sạch 710, và vùng không được làm mới/bản 712 trên FIG.7.

Trong quy trình tìm kiếm chuyển động 800, bộ mã hóa được hạn chế hoặc được ngăn không cho chọn vector chuyển động 812 bất kỳ mà dẫn tới một số trong số các mẫu của khối tham chiếu 814 đang nằm bên ngoài vùng được làm mới 806. Điều này xảy ra ngay cả khi khối tham chiếu 814 cung cấp các tiêu chuẩn về giá trị độ méo tốc độ tốt nhất khi dự đoán khối hiện tại 816 trong ảnh hiện tại 802. Do vậy, FIG.8 minh họa lý do cho khả năng không tối ưu trong chuyển động tìm kiếm 800 khi sử dụng sự hạn chế bộ mã hóa để hỗ trợ GDR.

FIG.9 minh họa một phương án thực hiện của dòng bit video 900. Như được sử dụng ở đây, dòng bit video 900 cũng có thể được gọi là dòng bit video được tạo mã, dòng bit, hoặc các biến thể của chúng. Như được thể hiện trên FIG.9, dòng bit 900 bao gồm ít nhất một đơn vị ảnh (PU: picture unit) 901. Mặc dù ba trong số các PU 901 được thể hiện trên FIG.9, số lượng PU 901 khác có thể có mặt trong dòng bit 900 trong các ứng dụng thực tế. Mỗi PU 901 là tập hợp của các cụm NAL mà được liên kết với nhau theo quy tắc phân loại được định rõ, là liên tiếp theo thứ tự giải mã, và chứa chính xác một ảnh đã được tạo mã (ví dụ, ảnh 914).

Theo một phương án, mỗi PU 901 chứa một hoặc nhiều trong số thành phần sau đây: thông tin công suất giải mã (DCI) 902, tập thông số video (VPS) 904, tập thông số chuỗi (SPS) 906, tập thông số ảnh (PPS) 908, tiêu đề ảnh (PH) 912, và ảnh 914. Mỗi một trong số DCI 902, VPS 904, SPS 906, và PPS 908 có thể được gọi chung là tập thông số. Theo một phương án, các tập thông số khác không được thể hiện trên FIG.9 cũng có thể có mặt trong dòng bit 900 như, ví dụ, tập thông số thích nghi (APS), vốn là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều lát cắt như được xác định bởi không hoặc nhiều phần tử cú pháp được tìm thấy trong các tiêu đề lát cắt.

DCI 902, mà cũng có thể được gọi là tập thông số giải mã (DPS) hoặc tập thông số bộ giải mã, là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho toàn bộ dòng bit. DCI 902 bao gồm các thông số mà giữ không đổi cho tuổi làm việc của dòng bit video (ví dụ, dòng bit 900), mà có thể biến đổi thành tuổi làm việc của phiên. DCI 902 có thể bao gồm tiêu sử, mức, và thông tin tiêu sử phụ để xác định điểm hoạt động liên hợp có độ phức tạp tối đa mà được bảo đảm không bao giờ bị vượt quá, ngay cả khi việc nối các chuỗi video xảy ra bên trong phiên. Còn bao gồm

theo cách tùy chọn các cờ ràng buộc, mà chỉ báo rằng dòng bit video sẽ được ràng buộc của việc sử dụng nhất định các đặc trưng như được chỉ báo bởi các giá trị của các cờ đó. Với điều này, dòng bit có thể được gắn nhãn là không sử dụng các công cụ nhất định, mà cho phép việc cấp phát tài nguyên khi thực hiện bộ giải mã trong số các vấn đề khác. Như tất cả các tập thông số, DCI 902 có mặt khi được tham chiếu đầu tiên, và được tham chiếu bởi ảnh đầu tiên nhất trong chuỗi video, ngụ ý rằng nó phải được gửi trong số các cụm NAL đầu tiên trong dòng bit. Mặc dù nhiều DCI 902 có thể nằm trong dòng bit, giá trị của các phần tử cú pháp trong đó không thể là không nhất quán khi được tham chiếu.

VPS 904 bao gồm sự phụ thuộc hoặc thông tin giải mã cho việc xây dựng tập ảnh tham chiếu của các lớp tăng cường. VPS 904 cung cấp hình phối cảnh hoặc hình vẽ toàn bộ của chuỗi có thể thay đổi được, bao gồm các loại nào của các điểm hoạt động được cung cấp, biên dạng, thứ hạng và mức của các điểm hoạt động, và một số thuộc tính mức cao khác của dòng bit mà có thể được sử dụng làm cơ sở cho sự thương lượng phiên và việc chọn nội dung, v.v.

SPS 906 chứa dữ liệu chung cho tất cả các ảnh trong chuỗi của các ảnh (SOP). SPS 906 là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều CLVS toàn cục như được xác định bởi nội dung của thành phần cú pháp được tìm thấy trong PPS được tham chiếu đến bởi thành phần cú pháp được tìm thấy trong mỗi tiêu đề ảnh. Ngược lại, PPS 908 chứa dữ liệu chung cho toàn bộ ảnh. PPS 908 là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho không hoặc nhiều ảnh toàn cục được tạo mã như được xác định bởi thành phần cú pháp được tìm thấy trong mỗi ảnh tiêu đề (ví dụ PH 912).

Theo một phương án, SPS 906 chứa cờ thứ nhất 930, vốn là cờ POC MSB được chỉ định là `sps_poc_msb_cycle_flag`. Khi cờ thứ nhất 930 có giá trị thứ nhất (ví dụ, 1), cờ thứ nhất 930 chỉ định rằng mỗi tiêu đề (ví dụ, tiêu đề ảnh 912) chứa cờ thứ hai 940, vốn là cờ POC MSB khác được chỉ định là `ph_poc_msb_cycle_present_flag`. Khi cờ thứ nhất 930 có thứ hai giá trị (ví dụ, 0), cờ thứ nhất 930 chỉ định rằng các tiêu đề (ví dụ, tiêu đề ảnh 912) không chứa cờ thứ hai 940. Như được sử dụng ở đây, cờ thứ hai 940 có thể được gọi là thành phần cú pháp `ph_poc_msb_cycle_present_flag`.

SPS 906 cũng chứa bộ chỉ báo của chiều dài 950 của giá trị POC MSB khi giá trị POC MSB có mặt trong tiêu đề tham chiếu đến SPS 906. Theo một phương án, bộ chỉ báo của chiều dài 950 được chỉ định là `sps_poc_msb_cycle_len_minus1`. Theo một phương án, `sps_poc_msb_cycle_len_minus1`, cộng 1, chỉ định chiều dài, trong các bit, của giá trị POC MSB khi giá trị POC MSB có mặt trong tiêu đề tham chiếu đến SPS. Theo một phương án, giá trị của bộ chỉ báo nằm trong khoảng từ không (0) đến $32 - \text{sps_log2_max_pic_order_cnt_lsb_minus4} - 5$, tuyến không loại. `sps_log2_max_pic_order_cnt_lsb_minus4` chỉ định giá trị của biến thiên `MaxPicOrderCntLsb` (ví dụ, số đếm thứ tự ảnh LSB tối đa) được sử dụng trong quy trình giải mã cho số đếm thứ tự ảnh như sau: $\text{MaxPicOrderCntLsb} = 2^{(\text{sps_log2_max_pic_order_cnt_lsb_minus4} + 4)}$, trong đó giá trị của `sps_log2_max_pic_order_cnt_lsb_minus4` sẽ nằm trong khoảng từ 0 đến 12, tuyến không loại.

Cờ thứ hai 940 chỉ định xem giá trị POC MSB 960 có mặt trong tiêu đề hay không. Theo một phương án, giá trị POC MSB 960 có mặt cho các AU được trộn chứa cả các ảnh IRAP và không có IRAP như đã được mô tả ở đây. Khi cờ thứ hai 940 có giá trị thứ nhất (ví dụ, 1), cờ thứ hai 940 chỉ định rằng giá trị POC MSB 960 có mặt trong tiêu đề (ví dụ, tiêu đề ảnh 912). Khi cờ thứ hai 940 có thứ hai giá trị (ví dụ, 0), cờ thứ hai 940 chỉ định rằng giá trị POC MSB 960 không có mặt trong tiêu đề. Theo một phương án, giá trị POC MSB 960 có thể được gọi là thành phần cú pháp hoặc các phần tử giá trị POC MSB.

Giá trị POC MSB 960 chỉ định giá trị của chu kỳ POC MSB của ảnh hiện tại (ví dụ, ảnh 914). Theo một phương án, giá trị POC MSB 960 được chỉ định là `ph_poc_msb_cycle_val`. Do vậy, giá trị POC MSB 960 được báo hiệu tường minh trong dòng bit 900 thay vì được suy ra dựa trên ảnh trước đó. Theo một phương án, giá trị POC MSB được báo hiệu trong dòng bit được sử dụng để thu được giá trị POC. Theo một phương án, giá trị POC được xác định bằng cách móc nối giá trị POC MSB và POC LSB, mà cũng có thể được báo hiệu trong dòng bit 900.

Mỗi ảnh được tạo mã được liên kết với biến số đếm thứ tự ảnh, được biểu thị là `PicOrderCntVal`. `PicOrderCntVal` được suy ra như sau: $\text{PicOrderCntVal} =$

PicOrderCntMsb + ph_pic_order_cnt_lsb, trong đó PicOrderCntMsb là POC MSB và ph_pic_order_cnt_lsb là POC LSB.

Như sẽ được giải thích một cách đầy đủ hơn bên dưới, việc sử dụng các cờ nêu trên cho phép các ảnh IRAP không đồng bộ cắt qua các lớp. Việc xử lý POC đã được mô tả ở đây đảm bảo rằng các giá trị POC của tất cả các ảnh bên trong AU là giống nhau. Để thực hiện điều này, các giá trị POC MSB được báo hiệu (thay vì được suy ra) cho các ảnh IRAP, tức là, các ảnh IDR, CRA và GDR. Điều này cho phép thiết kế đơn giản hơn đáng kể đối với việc suy ra POC so với việc thiết lập lại POC hỗn hợp và cơ chế báo hiệu POC MSB trong SHVC và MV-HEVC.

DCI 902, VPS 904, SPS 906, và PPS 908 có mặt trong các loại khác nhau của các cụm lớp trừu tượng mạng (NAL: network abstraction layer). Cụm NAL là cấu trúc cú pháp chứa chỉ báo về loại của dữ liệu cần tuân theo (ví dụ, dữ liệu video được tạo mã). Các cụm NAL được phân loại thành lớp tạo mã video (VCL) và các cụm không phải VCL NAL. Các cụm VCL NAL chứa dữ liệu mà thể hiện các giá trị của các mẫu trong các ảnh video, và các cụm không phải VCL NAL chứa thông tin bổ sung kết hợp bất kỳ như các tập thông số (dữ liệu quan trọng mà có thể áp dụng cho số lượng các cụm VCL NAL) và thông tin tăng cường bổ sung (thông tin định thời và dữ liệu bổ sung khác mà có thể tăng cường khả năng sử dụng của tín hiệu video được giải mã nhưng là không cần thiết cho việc giải mã các giá trị của các mẫu trong các ảnh video).

Theo một phương án, DCI 902 có mặt trong cụm không phải VCL NAL được chỉ định làm cụm DCI NAL hoặc cụm DPS NAL. Tức là, cụm DCI NAL có kiểu cụm DCI NAL (NUT) và cụm DPS NAL có DPS NUT. Theo một phương án, VPS 904 có mặt trong cụm không phải VCL NAL được chỉ định làm cụm VPS NAL. Do đó, cụm VPS NAL có VPS NUT. Theo một phương án, SPS 906 là cụm không phải VCL NAL được chỉ định làm cụm SPS NAL. Do đó, cụm SPS NAL có SPS NUT. Theo một phương án, PPS 908 có mặt trong cụm không phải VCL NAL được chỉ định làm cụm PPS NAL. Do đó, cụm PPS NAL có PPS NUT.

PH 912 là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho tất cả các lát cắt (ví dụ, các lát cắt 918) của ảnh đã được tạo mã (ví dụ, ảnh 914). Theo một phương án, PH 912 nằm trong kiểu mới của cụm không phải VCL NAL được chỉ định cụm PH NAL. Do đó, cụm PH NAL có PH NUT (ví dụ, PH_NUT). Theo một phương

án, có một và chỉ một PH 912 được chứa trong mỗi PU 901. Tức là, PU 901 chứa một hoặc duy nhất PH 912. Theo một phương án, chính xác là một cụm PH NAL có mặt cho mỗi ảnh 914 trong dòng bit 900.

Theo một phương án, cụm PH NAL được liên kết với PH 912 có ID thời gian và lớp ID. ID thời gian chỉ báo vị trí của cụm PH NAL, theo thời gian, tương đối với các cụm PH NAL khác trong dòng bit (ví dụ, dòng bit 900). ID lớp chỉ báo lớp (ví dụ, lớp 531 hoặc lớp 532) mà chứa cụm PH NAL. Theo một phương án, ID thời gian tương tự với, nhưng khác với, POC. POC chỉ xác định mỗi ảnh theo thứ tự. Trong một lớp dòng bit, ID thời gian và POC sẽ là giống nhau. Trong dòng bit đa lớp (ví dụ, xem FIG.5), các ảnh trong AU giống nhau sẽ có các POC khác nhau, nhưng có ID thời gian giống nhau.

Theo một phương án, cụm PH NAL ở trước cụm VCL NAL chứa lát cắt thứ nhất 918 của ảnh kết hợp 914. Điều này thiết lập sự liên kết giữa PH 912 và các lát cắt 918 của ảnh 914 được liên kết với PH 912 mà không cần có ID tiêu đề ảnh được báo hiệu trong PH 912 và được tham chiếu đến từ tiêu đề lát 920. Do đó, có thể suy ra rằng tất cả các cụm VCL NAL giữa hai PH 912 thuộc về ảnh giống nhau 914 và rằng ảnh 914 được liên kết với PH thứ nhất 912 trong số hai PH 912. Theo một phương án, cụm VCL NAL thứ nhất mà tuân theo PH 912 chứa lát cắt thứ nhất 918 của ảnh 914 được liên kết với PH 912.

Theo một phương án, cụm PH NAL tuân theo các tập thông số mức ảnh (ví dụ, PPS) hoặc các tập thông số mức cao hơn như DCI (tức là, DPS), VPS, SPS, PPS, v.v. có cả ID thời gian và ID lớp nhỏ hơn ID thời gian và ID lớp của cụm PH NAL, một cách tương ứng. Do đó, các tập thông số đó không được lặp lại bên trong ảnh hoặc cụm truy nhập. Do thứ tự này, PH 912 có thể được giải quyết ngay. Tức là, các tập thông số mà chứa các thông số thích hợp cho toàn bộ ảnh được định vị trong dòng bit trước cụm PH NAL. Không tập nào mà chứa các thông số cho một phần của ảnh được định vị sau cụm PH NAL.

Trong một phương án khác, cụm PH NAL tuân theo các tập thông số mức ảnh và các thông điệp thông tin tăng cường bổ sung tiền tố (SEI), hoặc các tập thông số mức cao hơn như thông điệp DCI (tức là, DPS), VPS, SPS, PPS, APS, SEI, v.v.

Theo một phương án, PH 912 có thể chứa thành phần cú pháp (ví dụ, cờ, v.v.) mà chỉ định kiểu ảnh của ảnh 914 được liên kết với PH 912. Kiểu ảnh có thể bao gồm, nhưng không bị giới hạn ở, các loại sau đây: ảnh nạp lại bộ giải mã tức thời (IDR), ảnh truy nhập ngẫu nhiên sạch (CRA), ảnh GDR, ảnh không có IRAP vốn là ảnh không phải GDR và chỉ chứa các lát cắt được dự đoán trong ảnh (I-slices), ảnh không có IRAP vốn là ảnh không phải GDR và chỉ chứa các lát cắt được dự đoán liên ảnh một chiều (P-slices) và các lát cắt I, và ảnh không có IRAP vốn là ảnh không phải GDR và chỉ chứa các lát cắt được dự đoán liên ảnh hai chiều (B-slices), các lát cắt P, và các lát cắt I. Vì vậy, một cờ trong PH 912 có thể chỉ báo xem tất cả các lát cắt (ví dụ, các lát cắt 918) trong ảnh (ví dụ, ảnh 914) là, ví dụ, các lát cắt của ảnh GDR (ví dụ, ảnh GDR 702). Điều này cũng hỗ trợ báo hiệu của số đếm thứ tự ảnh điểm hồi phục (POC) cho ảnh GDR ngay trong PH 912 thay vì trong mỗi tiêu đề lát (ví dụ, tiêu đề lát 920).

Theo một phương án, một hoặc nhiều phần tử cú pháp được báo hiệu trong PH 912 thay vì các tiêu đề lát cắt 920 của ảnh 914 được liên kết với PH 912. Các phần tử cú pháp đó là PPS ID được tham chiếu đến bởi ảnh 914, cờ chỉ định xem ảnh 914 là ảnh tham chiếu hay không, màu sắc mặt phẳng của ảnh 914, bit có nghĩa tối thiểu POC (LSB) của ảnh 914, điểm hồi phục POC nếu ảnh 914 là ảnh GDR (ví dụ, ảnh GDR 702), cờ chỉ định xem các ảnh phía trước khi ảnh 914 được xuất ra, và cờ chỉ định xem ảnh 914 là ảnh đầu ra hoặc không. PPS ID là bộ nhận dạng để xác định PPS cụ thể cho ảnh 914. Màu sắc mặt phẳng của ảnh 914 bao gồm luma và các thành phần chroma (ví dụ, Y, Cb, và Cr, v.v.) cho ảnh 914. POC LSB là bit (hoặc các bit) để xác định POC. POC là biến mà được liên kết với mỗi ảnh (ví dụ, ảnh 914), chỉ xác định ảnh kết hợp trong số các tất cả các ảnh trong CLVS, và, khi ảnh kết hợp sẽ được xuất ra từ DPB, chỉ báo vị trí của ảnh kết hợp theo thứ tự xuất tương đối với các vị trí thứ tự đầu ra của các ảnh khác trong CLVS giống nhau mà sẽ được xuất ra từ DPB. Điểm hồi phục POC là POC để xác định điểm hồi phục và do đó, xác định ảnh điểm hồi phục.

Chuyển các phần tử cú pháp này từ các tiêu đề lát cắt 920 từ PH 912 cho phép các phần tử cú pháp sẽ được báo hiệu ngay cho toàn bộ ảnh 914 thay vì lặp lại các phần tử cú pháp trong mỗi tiêu đề lát 920. Điều này làm giảm sự dự phòng và tăng hiệu suất tạo mã.

Ảnh 914 là mảng của các mẫu luma trong định dạng đơn sắc hoặc mảng của các mẫu luma và hai mảng tương ứng của các mẫu chroma trong định dạng màu 4:2:0, 4:2:2, và 4:4:4. Theo một phương án, có một và chỉ một ảnh 914 được chứa trong mỗi PU 901. Như vậy, có chỉ một PH 912 và chỉ một ảnh 914 tương ứng với PH 912 đó trong mỗi PU 901. Tức là, PU 901 chứa một hoặc duy nhất ảnh 914.

Ảnh 914 có thể là hoặc khung hoặc trường. Tuy nhiên, trong một CVS 916, hoặc tất cả các ảnh 914 là các khung hoặc tất cả các ảnh 914 là các trường. CVS 916 là chuỗi video đã được tạo mã cho mỗi chuỗi video lớp được tạo mã (CLVS) trong dòng bit video 900. Đáng chú ý là, CVS 916 và CLVS là giống nhau khi dòng bit video 900 bao gồm một lớp. CVS 916 và CLVS chỉ khác nhau khi dòng bit video 900 bao gồm nhiều lớp (ví dụ, như được thể hiện trên Fig.5).

Nói chung các PU 901 có thể bao gồm CLVS. CLVS là chuỗi của các PU 901 với giá trị giống nhau của `nuh_layer_id` mà bao gồm, theo thứ tự giải mã, PU bắt đầu chuỗi video lớp được tạo mã (CLVSS: coded layer video sequence start), theo sau bởi không hoặc nhiều PU 901 mà không là các PU CLVSS, bao gồm tất cả tiếp theo các PU 901 lên tới nhưng không bao gồm bất kỳ tiếp theo PU 901 là PU CLVSS. PU CLVSS là PU 901 trong đó ảnh đã được tạo mã (ví dụ, ảnh 914) là ảnh LVSS. Ảnh LVSS là ảnh đã được tạo mã mà là ảnh IRAP với `NoOutputBeforeRecoveryFlag` bằng 1 hoặc ảnh GDR (ví dụ, ảnh GDR 702) với `NoOutputBeforeRecoveryFlag` bằng 1.

Mỗi ảnh 914 chứa một hoặc nhiều lát cắt 918. Lát cắt 918 là số nguyên của các ảnh xếp lớp hoàn thành hoặc số nguyên của các hàng CTU hoàn thành liên tiếp bên trong ảnh xếp lớp của ảnh (ví dụ, ảnh 914). Mỗi lát cắt 918 được chứa riêng trong một cụm NAL (ví dụ, cụm VCL NAL). Ảnh xếp lớp (không được thể hiện trên hình vẽ) là vùng hình chữ nhật của các CTU bên trong cột ảnh xếp lớp riêng biệt và hàng ảnh xếp lớp riêng biệt trong ảnh (ví dụ ảnh 914). CTU (không được thể hiện trên hình vẽ) là CTB của các mẫu độ chói, hai CTB tương ứng của các mẫu sắc độ của hình ảnh có ba mảng mẫu, hoặc CTB của các mẫu của ảnh đen trắng hoặc ảnh mà được tạo mã sử dụng ba mặt phẳng màu riêng biệt và các cấu trúc cú pháp được sử dụng để tạo mã các mẫu. CTB (không được thể hiện trên hình vẽ) là khối NxN của mẫu đối với một số giá trị của N sao cho sự phân chia của thành phần thành CTB là một lần phân chia.

Khôi (không được thể hiện trên hình vẽ) là mảng $M \times N$ (M-cột x N-hàng) của các mẫu (ví dụ các điểm ảnh), hoặc mảng $M \times N$ của các hệ số biến đổi.

Theo một phương án, mỗi lát cắt 918 chứa lát cắt tiêu đề 920. Lát cắt tiêu đề 920 là một phần của lát cắt được tạo mã 918 chứa các phần tử dữ liệu thuộc về tất cả các ảnh xếp lớp hoặc các hàng CTU bên trong ảnh xếp lớp được biểu diễn trong lát cắt 918. Tức là, tiêu đề lát 920 chứa thông tin về lát cắt 918 chẳng hạn như loại lát cắt, mà các ảnh tham chiếu của chúng sẽ được sử dụng, v.v..

Các ảnh 914 và các lát cắt 918 của chúng bao gồm bao gồm dữ liệu được liên kết với các hình ảnh hoặc video được mã hóa hoặc được giải mã. Vì vậy, các ảnh 914 và các lát cắt 918 của chúng có thể được gọi một cách đơn giản là phần tải dữ liệu hoặc dữ liệu được mang trong dòng bit 900.

Người có hiểu biết trung bình trong lĩnh vực tương ứng sẽ hiểu rõ rằng dòng bit 900 có thể chứa các thông số và thông tin khác trong các ứng dụng thực tế.

FIG.10 là một phương án thực hiện của phương pháp 1000 để giải mã dòng bit video được tạo mã mà được thực hiện bởi bộ giải mã video (ví dụ, bộ giải mã video 400). Phương pháp 1000 có thể được thực hiện sau khi dòng bit được nhận trực tiếp hoặc gián tiếp từ bộ mã hóa video (ví dụ, bộ mã hóa video 300). Phương pháp 1000 cải thiện quy trình giải mã bằng cách cung cấp các công nghệ mà đơn giản hóa sự suy ra POC trong tạo mã video đa lớp khi AU giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị POC MSB được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo

cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Trong khối 1002, bộ giải mã video nhận dòng bit bao gồm SPS (ví dụ, SPS 906) chứa cờ thứ nhất (ví dụ, cờ thứ nhất 930) và bộ chỉ báo của chiều dài (ví dụ, bộ chỉ báo của chiều dài 950) của giá trị POC MSB (ví dụ, giá trị POC MSB 960) khi giá trị POC MSB có mặt trong tiêu đề (ví dụ, tiêu đề ảnh 912) tham chiếu đến SPS. Theo một phương án, cờ thứ nhất là cờ POC MSB được chỉ định là `sps_poc_msb_cycle_flag`. Giá trị của cờ thứ nhất chỉ định xem cờ thứ hai (ví dụ, cờ thứ hai 940) có mặt trong tiêu đề tham chiếu đến SPS hay không. Theo một phương án, khi giá trị của cờ thứ nhất là 1, cờ thứ hai có mặt trong mỗi tiêu đề tham chiếu đến SPS. Theo một phương án, khi giá trị của cờ thứ nhất là 0, cờ thứ hai không có mặt trong các tiêu đề tham chiếu đến SPS.

Theo một phương án, cờ thứ hai là cờ POC MSB khác và được chỉ định là `ph_poc_msb_cycle_present_flag`. Giá trị của cờ thứ hai chỉ định xem giá trị POC MSB có mặt trong tiêu đề hay không. Theo một phương án, giá trị POC MSB 960 có mặt cho các AU được trộn chứa cả các ảnh IRAP và không có IRAP như đã được mô tả ở đây. Theo một phương án, khi giá trị của cờ thứ hai là 1, giá trị POC MSB có mặt trong tiêu đề tham chiếu đến SPS. Theo một phương án, khi giá trị của cờ thứ hai là 0, giá trị POC MSB không có mặt trong tiêu đề tham chiếu đến SPS.

Trong khối 1004, bộ giải mã video xác định giá trị POC dựa trên giá trị POC MSB. Giá trị POC là biến được liên kết với mỗi ảnh. Giá trị POC chỉ xác định ảnh kết hợp trong số các tất cả các ảnh trong CLVS, xác định khi ảnh kết hợp sẽ được xuất ra từ DPB, và chỉ báo vị trí của ảnh kết hợp theo thứ tự xuất tương đối với các vị trí thứ tự đầu ra của các ảnh khác trong CLVS giống nhau mà sẽ được xuất ra từ DPB. Theo một phương án, giá trị POC được xác định bằng cách móc nối hoặc theo cách khác kết hợp giá trị POC MSB và giá trị LSB POS (ví dụ, giá trị $POC = POC\ MSB + LSB\ POC$).

Trong khối 1006, bộ giải mã video xác định ảnh (ví dụ, ảnh 914) từ dòng bit dựa trên giá trị POC. Trong khối 1008, bộ giải mã video giải mã ảnh để thu được ảnh đã được giải mã. Ngay khi được giải mã, ảnh có thể được sử dụng để tạo ra hoặc tạo ra hình ảnh hoặc chuỗi video để hiển thị cho người dùng trên màn hiển thị hoặc màn hình

của thiết bị điện tử (ví dụ, điện thoại "thông minh", máy tính bảng, máy tính xách tay, máy tính cá nhân, v.v.).

FIG.11 thể hiện một phương án thực hiện của phương pháp 1100 để mã hóa dòng bit video mà được thực hiện bởi bộ mã hóa video (ví dụ, bộ mã hóa video 300). Phương pháp 1100 có thể được thực hiện khi ảnh (ví dụ, từ video) sẽ được mã hóa thành dòng bit video và sau đó được truyền tới bộ giải mã video (ví dụ, bộ giải mã video 400). Phương pháp 1100 cải thiện quy trình mã hóa bằng cách cung cấp các công nghệ mà đơn giản hóa sự suy ra POC trong tạo mã video đa lớp khi AU giống nhau chứa cả các ảnh IRAP và không có IRAP. Trong một lớp tạo mã video, giá trị POC MSB được thiết lập lại khi gặp phải ảnh IRAP. Trong tạo mã video đa lớp trong đó AU giống nhau chứa cả các ảnh IRAP và không có IRAP trong các lớp khác nhau, thiết lập lại giá trị POC MSB khi gặp phải ảnh IRAP sẽ dẫn tới ảnh IRAP và ảnh không có IRAP có các giá trị POC MSB khác nhau, mà có thể gây ra các lỗi tạo mã. Để giải quyết điều này, thay vì thiết lập lại giá trị POC MSB cho ảnh IRAP, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được thu nhận từ dòng bit. Tức là, giá trị POC MSB cho ảnh IRAP và chiều dài POC MSB cho ảnh IRAP được báo hiệu trong dòng bit. Điều này đảm bảo rằng các giá trị POC MSB của các ảnh IRAP và không có IRAP trong AU giống nhau vẫn giống nhau, mà loại bỏ được các lỗi tạo mã tiềm tàng. Do vậy, bộ tạo mã/bộ giải mã (tức là, “phương tiện mã hóa”) trong tạo mã video được cải thiện tương đối với các codec hiện tại. Đối với vấn đề thực hành, quy trình tạo mã video được cải thiện tạo cho người dùng trải nghiệm người dùng tốt hơn khi các video được gửi, nhận và/hoặc xem.

Trong khối 1102, bộ mã hóa video thiết lập cờ thứ nhất (ví dụ, cờ thứ nhất 930) trong SPS (ví dụ, SPS 906) tới giá trị để chỉ định rằng cờ thứ hai (ví dụ, cờ thứ hai 940) có mặt trong tiêu đề tham chiếu đến SPS. Theo một phương án, cờ thứ nhất là cờ POC MSB được chỉ định là `sps_poc_msb_cycle_flag`. Theo một phương án, khi giá trị của cờ thứ nhất là 1, cờ thứ hai có mặt trong tiêu đề tham chiếu đến SPS. Theo một phương án, khi giá trị của cờ thứ nhất là 0, cờ thứ hai không có mặt trong tiêu đề tham chiếu đến SPS.

Trong khối 1104, bộ mã hóa video cung cấp bộ chỉ báo của chiều dài (ví dụ, bộ chỉ báo của chiều dài 950) của giá trị POC MSB (ví dụ, giá trị POC MSB 960) trong SPS khi giá trị POC MSB có mặt trong tiêu đề tham chiếu đến SPS.

Trong khối 1106, bộ mã hóa video thiết lập cờ thứ hai tới giá trị để chỉ định rằng giá trị POC MSB có mặt trong tiêu đề. Theo một phương án, cờ thứ hai là cờ POC MSB khác và được chỉ định là `ph_poc_msb_cycle_present_flag`. Theo một phương án, khi giá trị của cờ thứ hai là 1, giá trị POC MSB có mặt trong tiêu đề tham chiếu đến SPS. Theo một phương án, khi giá trị của cờ thứ hai là 0, giá trị POC MSB không có mặt trong tiêu đề tham chiếu đến SPS.

Trong khối 1108, bộ mã hóa video mã hóa SPS và tiêu đề vào trong dòng bit. Trong khối 1110, bộ mã hóa video lưu trữ dòng bit để truyền thông về phía bộ giải mã video. Dòng bit video có thể được lưu trữ trong bộ nhớ cho đến khi dòng bit video được truyền về phía bộ giải mã video. Ngay khi được nhận bởi bộ giải mã video, dòng bit video được mã hóa có thể được giải mã (ví dụ, như được mô tả trên đây) để tạo ra hoặc sản xuất hình ảnh hoặc chuỗi video để hiển thị cho người dùng trên màn hiển thị hoặc màn hình của thiết bị điện tử (ví dụ, điện thoại thông minh, máy tính bảng, máy tính xách tay, máy tính cá nhân, v.v.).

Cú pháp và các ngữ nghĩa sau đây có thể được sử dụng để thực hiện phương án được bộc lộ ở đây. Phần mô tả sau đây là tương đối với văn bản cơ bản, vốn là đặc tả kỹ thuật của bản thảo VVC cuối cùng.

Theo một phương án, ảnh kết hợp IRAP (của ảnh cụ thể với `nuh_layer_id` bằng giá trị cụ thể `layerId`) là ảnh trước đó IRAP theo thứ tự giải mã với `nuh_layer_id` bằng `layerId` (khi có mặt) mà không có ảnh GDR với `nuh_layer_id` bằng `layerId` giữa chúng. `nuh_layer_id` chỉ định bộ nhận dạng của lớp mà cụm VCL NAL thuộc về nó hoặc bộ nhận dạng của lớp mà cụm không phải VCL NAL áp dụng vào đó. Giá trị của `nuh_layer_id` sẽ nằm trong khoảng từ 0 đến 55, tuyến không loại. Các giá trị khác cho `nuh_layer_id` được dành riêng cho việc sử dụng trong tương lai bởi ITU-T | ISO/IEC.

Như được sử dụng ở đây, PU cũng có thể được gọi là cụm truy nhập lớp. Theo một phương án, cụm truy nhập lớp truy nhập ngẫu nhiên sạch (CRA) là cụm truy nhập lớp trong đó ảnh đã được tạo mã là ảnh CRA. Theo một phương án, ảnh truy nhập ngẫu nhiên sạch (CRA) là ảnh IRAP mà đối với nó mỗi cụm VCL NAL có

nal_unit_type bằng CRA_NUT. Theo một phương án, ảnh CRA không sử dụng dự đoán liên ảnh trong quy trình giải mã của nó, và có thể là ảnh thứ nhất trong dòng bit theo thứ tự giải mã, hoặc có thể xuất hiện sau trong dòng bit. Theo một phương án, ảnh CRA có thể kết hợp các ảnh RADL hoặc RASL. Khi ảnh CRA có NoOutputBeforeRecoveryFlag bằng 1, các ảnh RASL được kết hợp không được xuất ra bởi bộ giải mã, do chúng có thể không giải mã được, do chúng có thể chứa các tham chiếu đến các ảnh mà không có mặt trong dòng bit.

Theo một phương án, chuỗi video lớp được tạo mã (CLVS) là chuỗi của các PU với giá trị giống nhau của nuh_layer_id mà bao gồm, theo thứ tự giải mã, PU CLVSS, theo sau bởi không hoặc nhiều PU vốn không là các PU CLVSS, bao gồm tất cả các PU tiếp theo lên tới nhưng không bao gồm PU tiếp theo bất kỳ vốn là PU CLVSS. Theo một phương án, PU CLVSS có thể là IDR PU, CRA PU, hoặc GDR PU. Giá trị của NoOutputBeforeRecoveryFlag bằng 1 cho mỗi IDR PU, và mỗi CRA PU mà có HandleCraAsClvsStartFlag bằng 1, và mỗi CRA hoặc GDR PU là PU thứ nhất trong lớp của dòng bit theo thứ tự giải mã hoặc PU thứ nhất trong lớp của dòng bit mà tuân theo cụm EOS NAL theo thứ tự giải mã.

Theo một phương án, cụm truy nhập lớp bắt đầu chuỗi video lớp được tạo mã (CLVSS) là cụm truy nhập lớp trong đó ảnh đã được tạo mã là ảnh LVSS. Ảnh bắt đầu chuỗi video lớp được tạo mã (CLVSS) là ảnh đã được tạo mã mà là ảnh IRAP với NoIncorrectPicOutputFlag bằng 1 hoặc ảnh GDR với NoIncorrectPicOutputFlag bằng 1.

Theo một phương án, chuỗi video đã được tạo mã (CVS: coded video sequence): Chuỗi của các cụm truy nhập mà bao gồm, theo thứ tự giải mã, cụm truy nhập CVSS, theo sau bởi không hoặc nhiều cụm truy nhập mà không là các cụm truy nhập CVSS, bao gồm tất cả các cụm truy nhập tiếp theo lên tới nhưng không bao gồm cụm truy nhập tiếp theo bất kỳ là cụm truy nhập CVSS.

Theo một phương án, cụm truy nhập bắt đầu chuỗi video đã được tạo mã (CVSS): Cụm truy nhập trong đó có cụm truy nhập lớp cho mỗi lớp trong CVS và ảnh đã được tạo mã trong mỗi cụm truy nhập lớp là ảnh LVSS. Chú ý rằng điều này yêu cầu rằng mỗi CVSS AU là AU hoàn thành, và mỗi ảnh trong CVSS AU là ảnh LVSS. Theo một phương án, cụm truy nhập làm mới việc giải mã dần dần (GDR): Cụm truy

nhập trong đó ảnh đã được tạo mã trong cụm truy nhập lớp hiện tại là ảnh GDR. Theo một phương án, cụm truy nhập lớp làm mới việc giải mã dần dần (GDR): Cụm truy nhập lớp trong đó ảnh đã được tạo mã là ảnh GDR. Theo một phương án, ảnh làm mới việc giải mã dần dần (GDR): Ảnh mà mỗi cụm VCL NAL cho nó có nal_unit_type bằng GDR_NUT.

Cú pháp SPS là như sau.

seq_parameter_set_rbsp() {	Bộ mô tả
...	
log2_max_pic_order_cnt_lsb_minus4	ue(v)
poc_msb_in_rap_pics_flag	u(1)
if(poc_msb_in_rap_pics_flag > 0)	
poc_msb_len_minus1	ue(v)
...	
}	

Theo một phương án, các phần tử cú pháp đậm nêu trên có thể có mặt trong VPS.

Tiêu đề lát cú pháp là như sau.

slice_header() {	Bộ mô tả
slice_pic_parameter_set_id	ue(v)
if(rect_slice_flag NumBricksInPic > 1)	
slice_address	u(v)
if(!rect_slice_flag && !single_brick_per_slice_flag)	
num_bricks_in_slice_minus1	ue(v)
non_reference_picture_flag	u(1)
slice_type	ue(v)
if(separate_colour_plane_flag == 1)	
colour_plane_id	u(2)
slice_pic_order_cnt_lsb	u(v)

if(poc_msb_in_rap_pics_flag && nal_unit_type >= IDR_W_RADL && nal_unit_type <= GDR_NUT)	
poc_msb_val	u(v)
if(nal_unit_type == GDR_NUT)	
recovery_poc_cnt	ue(v)
...	
}	

Theo một phương án, phần điều kiện cú pháp “&& nal_unit_type >= IDR_W_RADL && nal_unit_type <= GDR_NUT” có thể được loại bỏ. Tức là, báo hiệu poc_msb_val cho các lát cắt với giá trị bất kỳ của nal_unit_type khi poc_msb_in_rap_pics_flag bằng 1.

Theo một phương án, bổ sung một cờ mức ảnh cho điều kiện về sự có mặt của thành phần cú pháp poc_msb_val. Tức là, cho phép các bộ mã hóa quyết định hoặc báo hiệu hoặc không báo hiệu poc_msb_val riêng biệt cho mỗi lát cắt của ảnh IRAP hoặc ảnh GDR khi poc_msb_in_rap_pics_flag bằng 1.

Theo một phương án, áp dụng cả hai cách nêu trên. Theo một phương án, báo hiệu poc_msb_val nhờ sử dụng se(v) hoặc ue(v), và loại bỏ thành phần cú pháp poc_msb_len_minus1 ra khỏi cú pháp SPS.

Các ngữ nghĩa tiêu đề cụm NAL được cung cấp.

nuh_layer_id chỉ định bộ nhận dạng của lớp mà cụm VCL NAL thuộc về nó hoặc bộ nhận dạng của lớp mà cụm không phải VCL NAL áp dụng vào đó.

Giá trị của nuh_layer_id sẽ giống nhau cho tất cả các cụm VCL NAL của ảnh đã được tạo mã. Giá trị của nuh_layer_id của ảnh đã được tạo mã hoặc cụm truy nhập lớp là giá trị của nuh_layer_id của các cụm VCL NAL của ảnh đã được tạo mã hoặc cụm truy nhập lớp.

nal_unit_type chỉ định kiểu cụm NAL, tức là, kiểu của cấu trúc dữ liệu RBSP chứa trong cụm NAL.

Giá trị của nal_unit_type sẽ giống nhau cho tất cả các cụm NAL lát cắt được tạo mã của ảnh. Ảnh hoặc cụm truy nhập lớp được gọi là có kiểu cụm NAL giống như

các cụm NAL lát cắt được tạo mã của ảnh hoặc cụm truy nhập lớp. Lưu ý rằng hai đoạn mô tả nêu trên cho phép các cụm VCL NAL bên trong AU có các kiểu cụm NAL khác nhau, mà cho phép RAP hỗn hợp và các ảnh không phải RAP bên trong AU. Ảnh RAP là ảnh IDR, CAR hoặc GDR.

Giá trị của `nal_unit_type` sẽ là giống nhau cho các ảnh của cụm truy nhập CVSS.

Thứ tự của các cụm truy nhập và sự liên kết với các CVS được đề xuất.

Dòng bit phù hợp với tiêu chuẩn bao gồm một hoặc nhiều CVS. CVS bao gồm một hoặc nhiều cụm truy nhập. Cụm truy nhập thứ nhất của CVS là cụm truy nhập CVSS, trong đó mỗi cụm thể hiện cụm truy nhập lớp là cụm truy nhập lớp CLVSS, vốn là cụm truy nhập lớp IRAP với `NoIncorrectPicOutputFlag` bằng 1 hoặc cụm truy nhập lớp GDR với `NoIncorrectPicOutputFlag` bằng 1. Mỗi cụm truy nhập CVSS sẽ có ảnh trong mỗi một trong số các lớp thể hiện trong CVS. Có yêu cầu về sự hợp chuẩn dòng bit rằng, khi có mặt, mỗi cụm truy nhập lớp trong cụm truy nhập tiếp theo sau cụm truy nhập mà chứa kết thúc của cụm NAL tiếp theo hoặc kết thúc của cụm NAL dòng bit sẽ là cụm truy nhập lớp IRAP, mà có thể là cụm truy nhập lớp IDR hoặc cụm truy nhập lớp CRA, hoặc cụm truy nhập lớp GDR.

Các ngữ nghĩa SPS được cung cấp.

`log2_max_pic_order_cnt_lsb_minus4` chỉ định giá trị của biến thiên `MaxPicOrderCntLsb` mà được sử dụng trong quy trình giải mã cho số đếm thứ tự ảnh như sau:

$$\text{MaxPicOrderCntLsb} = 2(\text{log2_max_pic_order_cnt_lsb_minus4} + 4)$$

Giá trị của `log2_max_pic_order_cnt_lsb_minus4` sẽ nằm trong khoảng từ 0 đến 12, tuyến không loại.

`poc_msb_in_rap_pics_flag` bằng 1 chỉ định rằng `poc_msb_val` thành phần cú pháp được báo hiệu cho các ảnh tham chiếu đến SPS và các cụm VCL NAL mà đối với chúng có `nal_unit_type` nằm trong khoảng từ `IDR_W_RADL` đến `GDR_NUT`, tuyến không loại. `poc_msb_in_rap_pics_flag` bằng 0 chỉ định rằng `poc_msb_val` thành phần cú pháp không được báo hiệu cho các ảnh tham chiếu đến SPS.

`poc_msb_len_minus1` cộng 1 chỉ định chiều dài, trong các bit, của các phần tử cú pháp `poc_msb_val`, khi có mặt. Giá trị của `poc_msb_len_minus1` sẽ nằm trong khoảng từ 0 đến $32 - \log_2 \text{max_pic_order_cnt_lsb_minus4} - 5$, tuyến không loại.

Các ngữ nghĩa tiêu đề lát được cung cấp.

`slice_pic_order_cnt_lsb` chỉ định số đếm thứ tự ảnh modulo `MaxPicOrderCntLsb` cho ảnh hiện tại. Chiều dài của thành phần cú pháp `slice_pic_order_cnt_lsb` là $\log_2 \text{max_pic_order_cnt_lsb_minus4} + 4$ bit. Giá trị của `slice_pic_order_cnt_lsb` sẽ nằm trong khoảng từ 0 đến `MaxPicOrderCntLsb - 1`, tuyến không loại.

`poc_msb_val` chỉ định giá trị POC MSB của ảnh hiện tại. Chiều dài của thành phần cú pháp `poc_msb_val` là `poc_msb_len_minus1 + 1` bit.

Quy trình giải mã cho ảnh đã được tạo mã được đề xuất.

`PictureOutputFlag` được chọn như sau:

- nếu một trong số các điều kiện sau đây là đúng, `PictureOutputFlag` được đặt bằng 0.

- ảnh hiện tại là ảnh RASL và `NoIncorrectPicOutputFlag` của ảnh kết hợp IRAP bằng 1.

- `gdr_enabled_flag` bằng 1 và ảnh hiện tại là ảnh GDR với `NoIncorrectPicOutputFlag` bằng 1.

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

- ảnh hiện tại không thuộc về lớp đầu ra.

- theo cách khác, `PictureOutputFlag` được đặt bằng `pic_output_flag`.

Quy trình giải mã cho số đếm thứ tự ảnh được đề xuất.

Đầu ra của quy trình này là `PicOrderCntVal`, số đếm thứ tự ảnh của ảnh hiện tại.

Mỗi ảnh được tạo mã được liên kết với biến số đếm thứ tự ảnh, được biểu thị là `PicOrderCntVal`.

Khi `poc_msb_val` không có mặt đối với ảnh và ảnh hiện tại không là ảnh LVSS, các biến số `prevPicOrderCntLsb` và `prevPicOrderCntMsb` được suy ra như sau:

– Let prevTid0Pic là ảnh trước đó theo thứ tự giải mã mà có nuh_layer_id bằng nuh_layer_id của ảnh hiện tại và TemporalId bằng 0 và không là ảnh RASL hoặc RADL.

– biến prevPicOrderCntLsb được đặt bằng slice_pic_order_cnt_lsb của prevTid0Pic.

– biến prevPicOrderCntMsb được đặt bằng PicOrderCntMsb của prevTid0Pic.

Biến PicOrderCntMsb của ảnh hiện tại được suy ra như sau:

– nếu poc_msb_val có mặt đối với ảnh, PicOrderCntMsb được đặt bằng poc_msb_val * MaxPicOrderCntLsb.

– theo cách khác (poc_msb_val không có mặt đối với ảnh), nếu ảnh hiện tại là ảnh LVSS, PicOrderCntMsb được đặt bằng 0.

– ngược lại, PicOrderCntMsb được suy ra như sau:

if((slice_pic_order_cnt_lsb < prevPicOrderCntLsb) &&
 ((prevPicOrderCntLsb - slice_pic_order_cnt_lsb) >= (MaxPicOrderCntLsb / 2)))

PicOrderCntMsb = prevPicOrderCntMsb + MaxPicOrderCntLsb

(8-1)

else if((slice_pic_order_cnt_lsb > prevPicOrderCntLsb) &&

((slice_pic_order_cnt_lsb - prevPicOrderCntLsb) > (MaxPicOrderCntLsb / 2)))

PicOrderCntMsb = prevPicOrderCntMsb - MaxPicOrderCntLsb

else

PicOrderCntMsb = prevPicOrderCntMsb

PicOrderCntVal được suy ra như sau:

PicOrderCntVal = PicOrderCntMsb + slice_pic_order_cnt_lsb

(8-2)

Lưu ý – Tất cả các ảnh CLVSS mà poc_msb_val không có mặt đối với nó sẽ có PicOrderCntVal bằng slice_pic_order_cnt_lsb do đối với các ảnh PicOrderCntMsb đó được đặt bằng 0.

Giá trị của PicOrderCntVal sẽ nằm trong khoảng từ -231 đến $231 - 1$, tuyên không loại.

Trong một CVS, các trị số PicOrderCntVal cho hai ảnh được tạo mã bất kỳ với giá trị giống nhau của nuh_layer_id shall sẽ không giống nhau.

Tất cả các ảnh trong cụm truy nhập cụ thể bất kỳ sẽ có cùng giá trị của PicOrderCntVal.

Hàm PicOrderCnt(picX) được định rõ như sau:

$$\text{PicOrderCnt(picX)} = \text{PicOrderCntVal của ảnh picX} \quad (8-3)$$

Hàm DiffPicOrderCnt(picA, picB) được định rõ như sau:

$$\text{DiffPicOrderCnt(picA, picB)} = \text{PicOrderCnt(picA)} - \text{PicOrderCnt(picB)} \quad (8-4)$$

Dòng bit sẽ không chứa dữ liệu mà tạo ra các giá trị của DiffPicOrderCnt(picA, picB) được sử dụng trong quy trình giải mã vốn không nằm trong khoảng từ -215 đến $215 - 1$, tuyên không loại.

Lưu ý – Lấy X là ảnh hiện tại và Y và Z là hai ảnh khác trong cùng CVS, Y và Z được xem là nằm trong cùng hướng thứ tự đầu ra từ X khi cả DiffPicOrderCnt(X, Y) và DiffPicOrderCnt(X, Z) là dương hoặc cả hai là âm.

FIG.12 là hình vẽ dạng giản lược của thiết bị mã hóa video 1200 (ví dụ, bộ mã hóa video 300 hoặc bộ giải mã video 400) theo một phương án thực hiện của sáng chế. Thiết bị tạo mã video 1200 là thích hợp để thực hiện các phương án được bộc lộ như được mô tả ở đây. Thiết bị tạo mã video 1200 bao gồm các cổng vào 1210 và các cụm nhận (Rx) 1220 để nhận dữ liệu; bộ xử lý, cụm logic, hoặc bộ xử lý trung tâm (Central Processing Unit, viết tắt là CPU) 1230 để xử lý dữ liệu; các cụm truyền (Tx) 1240 và các cổng ra 1250 để truyền dữ liệu; và bộ nhớ 1260 để lưu trữ dữ liệu. Thiết bị tạo mã video 1200 cũng có thể bao gồm các thành phần quang-sang-điện (Optical-to-Electrical, viết tắt là OE) và các thành phần điện sang quang (Electrical-to-Optical, viết tắt là EO) được ghép nối với các cổng vào 1210, các cụm nhận 1220, các cụm truyền 1240, và các cổng ra 1250 để thu vào hoặc đưa ra các tín hiệu quang hoặc điện.

Bộ xử lý 1230 được thực hiện bởi phần cứng và phần mềm. Bộ xử lý 1230 có thể được thực hiện như là một hoặc nhiều chip CPU, các lõi (chẳng hạn, như bộ xử lý đa lõi), các mảng cổng khả lập trình bằng trường (field-programmable gate array, viết

tất là FPGA), các mạch tích hợp ứng dụng cụ thể (application specific integrated circuit, viết tắt là ASIC), và các bộ xử lý tín hiệu số (digital signal processor, viết tắt là DSP). Bộ xử lý 1230 truyền thông với các cổng vào 1210, các cụm nhận 1220, các cụm truyền 1240, các cổng ra 1250, và bộ nhớ 1260. Bộ xử lý 1230 bao gồm môđun tạo mã 1270. Môđun tạo mã 1270 thực hiện các phương án nêu trên. Ví dụ, môđun tạo mã 470 thực hiện, xử lý, chuẩn bị, hoặc cung cấp các chức năng mã hóa khác nhau. Do đó, việc bao gồm môđun tạo mã 1270 cung cấp sự cải tiến cơ bản đối với chức năng của thiết bị tạo mã video 1200 và tạo hiệu quả biến đổi thiết bị tạo mã video 1200 sang trạng thái khác nhau. Theo cách khác, môđun tạo mã 1270 được thực hiện dưới dạng các lệnh được lưu trữ trong bộ nhớ 1260 và được thực hiện bởi bộ xử lý 1230.

Thiết bị mã hóa video 1200 có thể còn bao gồm các thiết bị vào và/hoặc ra (I/O) 1280 để truyền thông dữ liệu đến và từ người dùng. Các thiết bị I/O 1280 có thể bao gồm các thiết bị đầu ra như màn hiển thị để hiển thị dữ liệu video, loa để xuất ra âm thanh dữ liệu, v.v. Các thiết bị I/O 1280 có thể còn bao gồm các thiết bị đầu vào, như bàn phím, chuột, bi xoay, v.v., và/hoặc các giao diện tương ứng để tương tác với các thiết bị đầu ra như vậy.

Bộ nhớ 1260 bao gồm một hoặc nhiều ổ đĩa, ổ băng, và ổ đĩa trạng thái rắn và có thể được sử dụng dưới dạng thiết bị lưu trữ dữ liệu tràn, để lưu trữ các chương trình khi các chương trình như vậy được chọn để thực hiện, và để lưu trữ các lệnh và dữ liệu được đọc trong suốt quá trình thực chương trình. Bộ nhớ 1260 có thể là khả biến và/hoặc bất khả biến và có thể là bộ nhớ chỉ đọc (ROM), bộ nhớ truy cập ngẫu nhiên (RAM), bộ nhớ định nội dung địa chỉ bậc ba (TCAM), và/hoặc bộ nhớ truy nhập ngẫu nhiên tĩnh (SRAM).

FIG.13 là hình vẽ dạng giản lược của một phương án thực hiện của phương tiện để tạo mã 1300. Theo một phương án, phương tiện để tạo mã 1300 được triển khai trong thiết bị mã hóa video 1302 (ví dụ, bộ mã hóa video 300 hoặc bộ giải mã video 400). Thiết bị mã hóa video 1302 bao gồm phương tiện nhận 1301. Phương tiện nhận 1301 được tạo cấu hình để nhận ảnh để mã hóa hoặc nhận dòng bit để giải mã. Thiết bị mã hóa video 1302 bao gồm phương tiện truyền 1307 được ghép nối với phương tiện nhận 1301. Phương tiện truyền 1307 được tạo cấu hình để truyền dòng bit tới bộ giải

mã hoặc truyền hình ảnh được giải mã tới phương tiện hiển thị (ví dụ, một trong số các thiết bị I/O 1280).

Thiết bị mã hóa video 1302 bao gồm phương tiện lưu trữ 1303. Phương tiện lưu trữ 1303 được lắp với ít nhất một trong số phương tiện nhận 1301 hoặc phương tiện truyền 1307. Phương tiện lưu trữ 1303 được tạo cấu hình để lưu trữ các lệnh. Thiết bị mã hóa video 1302 cũng bao gồm phương tiện xử lý 1305. Phương tiện xử lý 1305 được lắp với phương tiện lưu trữ 1303. Phương tiện xử lý 1305 được tạo cấu hình để thực thi các lệnh được lưu trữ trong phương tiện lưu trữ 1303 để thực hiện các phương pháp được bộc lộ ở đây.

Cũng nên hiểu được rằng các bước của các phương pháp làm ví dụ mà được đề cập đến ở đây không nhất thiết phải được thực hiện theo thứ tự được mô tả, và thứ tự của các bước của các phương pháp này nên được hiểu chỉ là ví dụ. Tương tự, các bước bổ sung có thể có trong các phương pháp này, và các bước cụ thể có thể được bỏ qua hoặc được kết hợp, trong các phương pháp thống nhất với các phương án thực hiện khác nhau của sáng chế.

Mặc dù một vài phương án đã được đề xuất theo sáng chế, nhưng nên hiểu được rằng các hệ thống và các phương pháp đã được bộc lộ có thể được thực hiện theo nhiều dạng cụ thể khác mà không trệtch khỏi nguyên lý hoặc phạm vi của sáng chế. Các ví dụ sẽ được xem là minh họa sáng chế và không làm giới hạn sáng chế, và mục đích không bị giới hạn ở các phần chi tiết nêu ra ở đây. Ví dụ, các phần tử hoặc các thành phần khác nhau có thể được kết hợp hoặc tích hợp trong hệ thống khác hoặc các dấu hiệu nhất định có thể được bỏ qua, hoặc không được thực hiện.

Ngoài ra, các kỹ thuật, các hệ thống, các hệ thống con, và các phương pháp được mô tả và được minh họa trong các phương án khác nhau là rời rạc hoặc tách biệt có thể được kết hợp hoặc được tích hợp với các hệ thống, các môđun, các kỹ thuật, hoặc các phương pháp khác mà không trệtch khỏi phạm vi của sáng chế. Các mục khác được thể hiện hoặc thảo luận như được ghép nối hoặc ghép nối trực tiếp hoặc giao tiếp với nhau có thể được ghép nối hoặc giao tiếp gián tiếp thông qua một số giao diện, thiết bị hoặc thành phần trung gian, cho dù bằng điện, cơ hay cách khác. Các ví dụ khác về các thay đổi, các thay thế, và các sửa đổi có thể thực hiện được bởi người có

hiểu biết trung bình trong lĩnh vực và có thể được tạo ra mà không trệch khỏi nguyên lý và phạm vi được bộc lộ ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã được thực hiện bởi bộ giải mã video bao gồm các bước:
nhận, bởi bộ giải mã video, dòng bit bao gồm tập thông số chuỗi (SPS) và tiêu đề tham chiếu đến SPS, trong đó SPS chứa cờ thứ nhất và bộ chỉ báo có độ dài của giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC), trong đó tiêu đề chứa cờ thứ hai, và giá trị POC MSB, trong đó giá trị của cờ thứ nhất chỉ định rằng cờ thứ hai có mặt trong tiêu đề, và trong đó giá trị của cờ thứ hai chỉ định rằng giá trị POC MSB có mặt trong tiêu đề; trong đó cờ thứ nhất là `sps_poc_msb_cycle_flag` và cờ thứ hai là `ph_poc_msb_cycle_present_flag`;
xác định, bởi bộ giải mã video, giá trị POC dựa trên giá trị POC MSB;
xác định, bởi bộ giải mã video, ảnh từ dòng bit dựa trên giá trị POC; và
giải mã, bởi bộ giải mã video, ảnh để thu được ảnh đã được giải mã.
2. Phương pháp theo điểm 1, trong đó giá trị của cờ thứ nhất là một.
3. Phương pháp theo điểm 1 hoặc 2, trong đó giá trị của cờ thứ hai là một.
4. Phương pháp theo điểm 1 hoặc 2, trong đó cờ thứ hai là cờ mức ảnh.
5. Phương pháp theo điểm 1 hoặc 2, trong đó bộ chỉ báo của chiều dài cộng 1 chỉ định chiều dài, trong các bit, của MSB POC.
6. Phương pháp mã hóa được thực hiện bởi bộ mã hóa video, phương pháp này bao gồm các bước:
thiết lập, bởi bộ mã hóa video, cờ thứ nhất trong tập thông số chuỗi (SPS) tới giá trị để chỉ định rằng cờ thứ hai có mặt trong tiêu đề tham chiếu đến SPS, trong đó cờ thứ nhất là `sps_poc_msb_cycle_flag` và cờ thứ hai là `ph_poc_msb_cycle_present_flag`;
cung cấp trong SPS, bởi bộ mã hóa video, bộ chỉ báo có độ dài của giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC) trong tiêu đề;

thiết lập, bởi bộ mã hóa video, cờ thứ hai trong tiêu đề tới giá trị để chỉ định rằng giá trị MSB POC có mặt trong tiêu đề;
mã hóa, bởi bộ mã hóa video, SPS và tiêu đề vào trong dòng bit; và
lưu trữ, bởi bộ mã hóa video, dòng bit cho việc truyền thông về phía bộ giải mã video.

7. Phương pháp theo điểm 6, trong đó giá trị của cờ thứ nhất là một.

8. Phương pháp theo điểm 6 hoặc 7, trong đó giá trị của cờ thứ hai là một.

9. Phương pháp theo điểm 6 hoặc 7, trong đó giá trị POC được xác định bằng bổ sung giá trị MSB POC cho giá trị các bit ít đáng kể nhất (LSB: least significant bit) POC, và trong đó cờ thứ hai là cờ mức ảnh.

10. Thiết bị giải mã, bao gồm:

bộ nhận được tạo cấu hình để nhận dòng bit bao gồm tập thông số chuỗi (SPS: sequence parameter set) chứa cờ thứ nhất, và khi giá trị các bit đáng kể nhất (MSB) số đếm thứ tự ảnh (POC) có mặt trong tiêu đề tham chiếu đến SPS, SPS bao gồm bộ chỉ báo về chiều dài của giá trị POC MSB;

bộ nhớ được ghép nối với bộ nhận, bộ nhớ để lưu trữ các lệnh; và

bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực thi các lệnh khiến cho thiết bị giải mã thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 5.

11. Thiết bị mã hóa bao gồm:

bộ nhớ chứa các lệnh;

bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh khiến cho thiết bị mã hóa thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 6 đến 9.

1/13

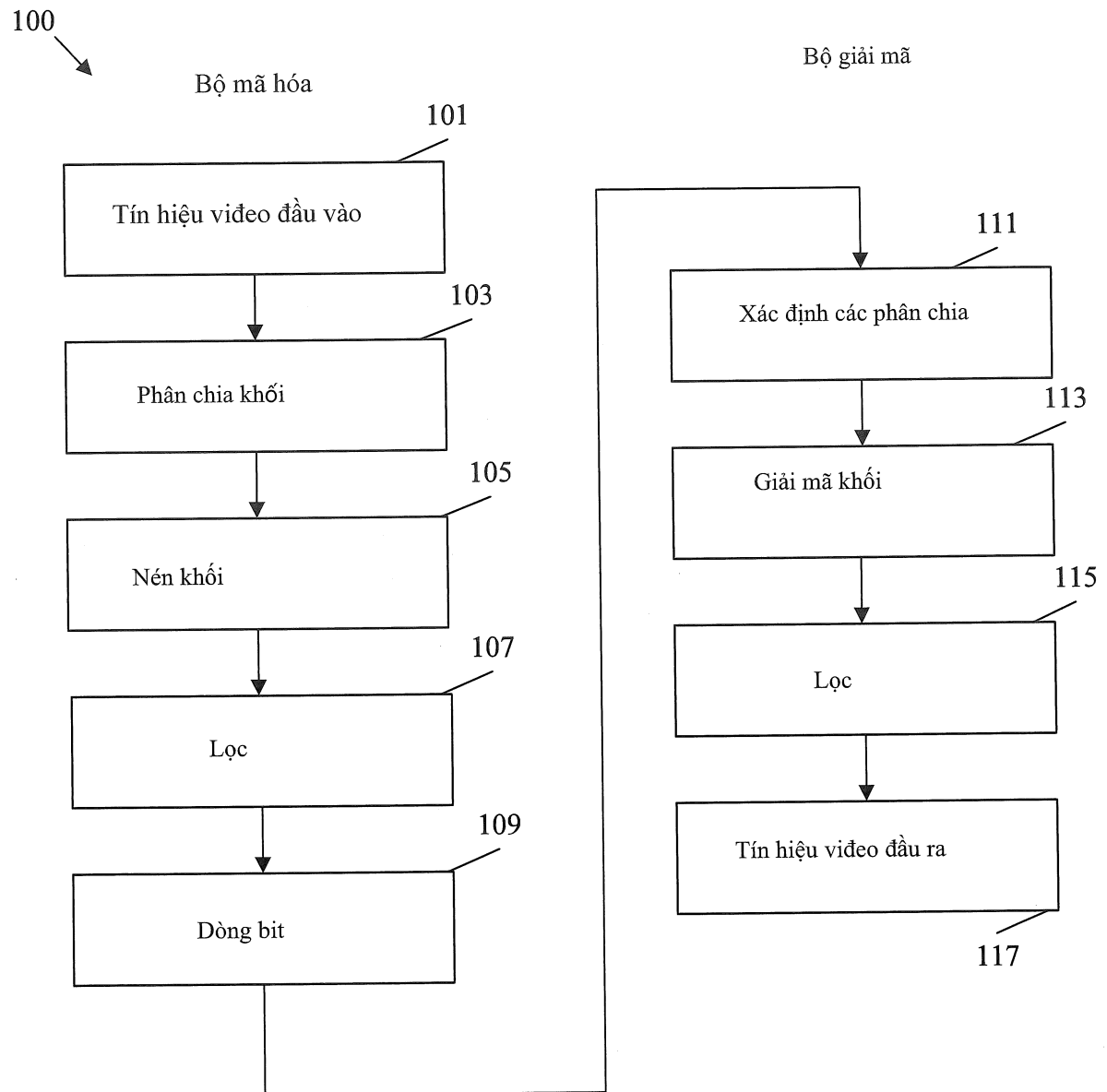


FIG. 1

2/13

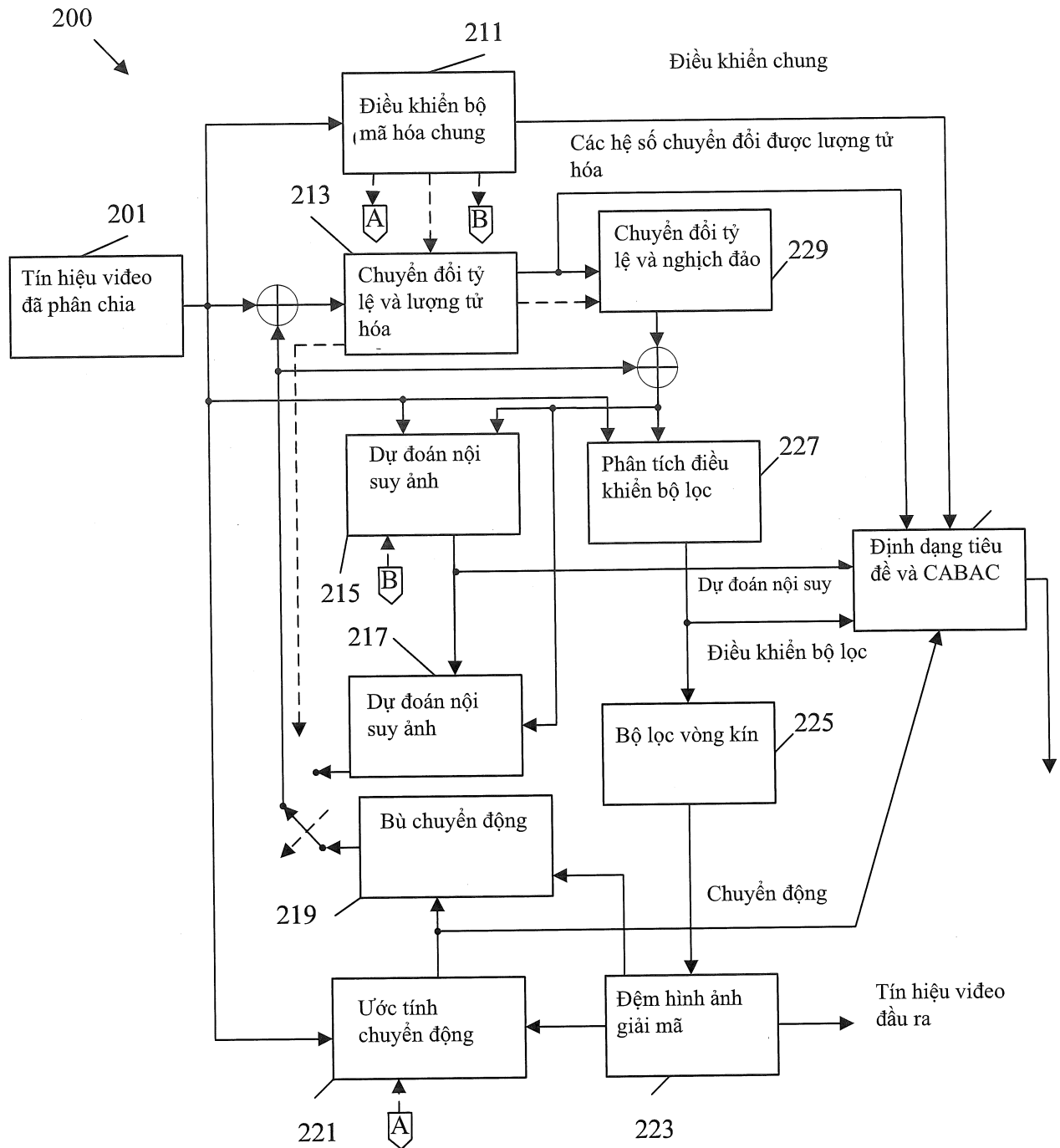


FIG. 2

3/13

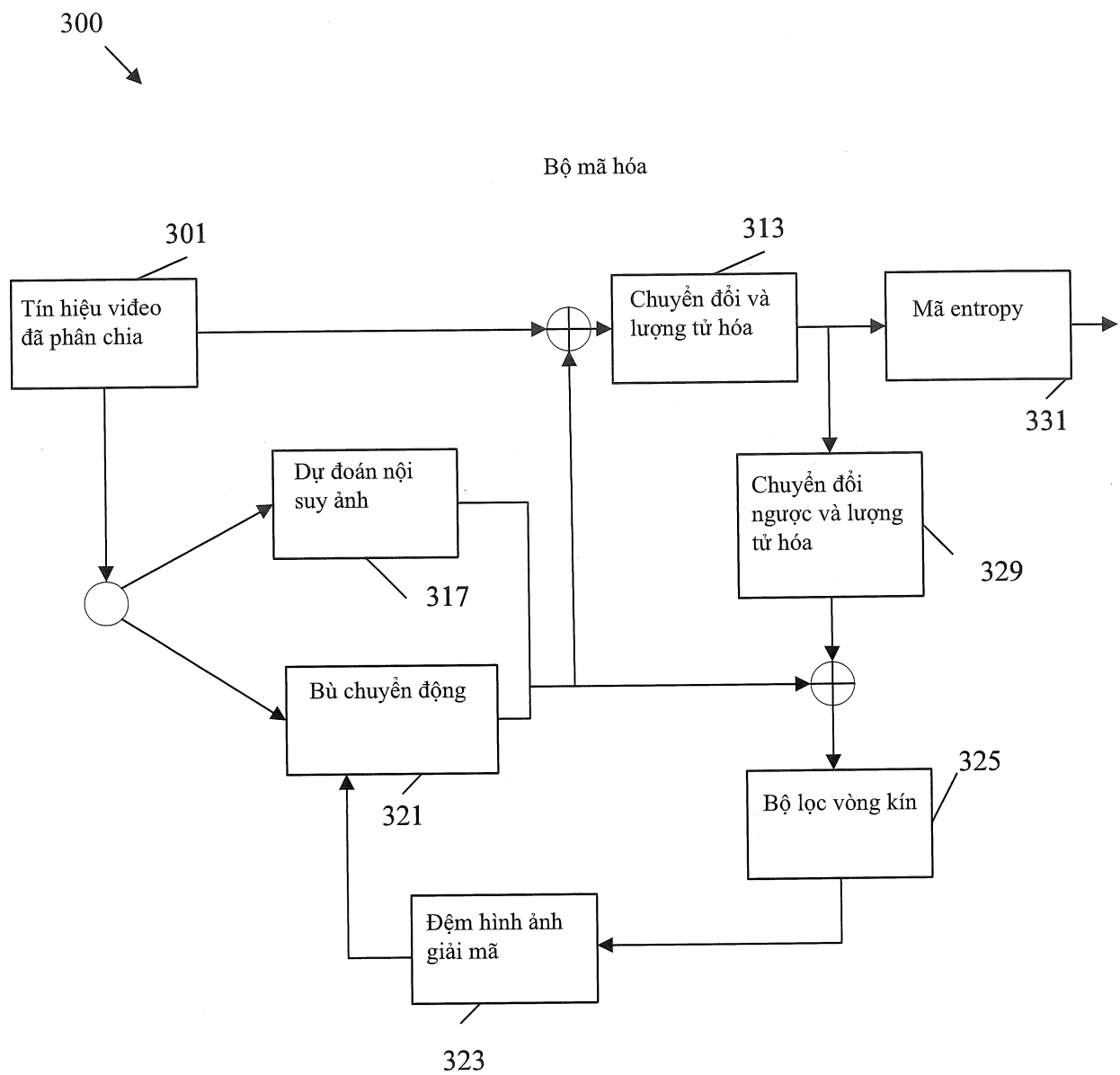


FIG. 3

4/13

400
↓

Bộ giải mã

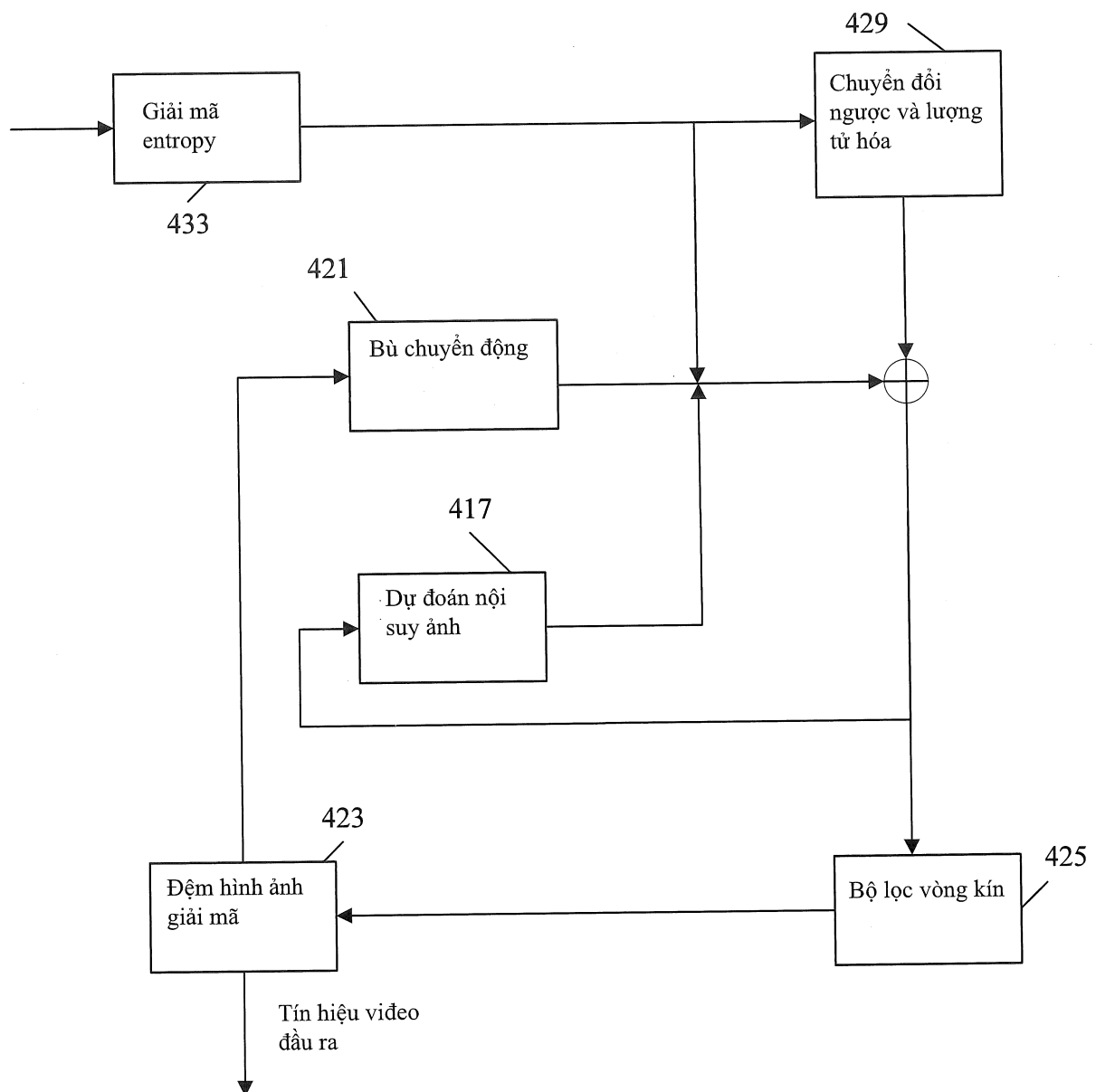


FIG. 4

5/13

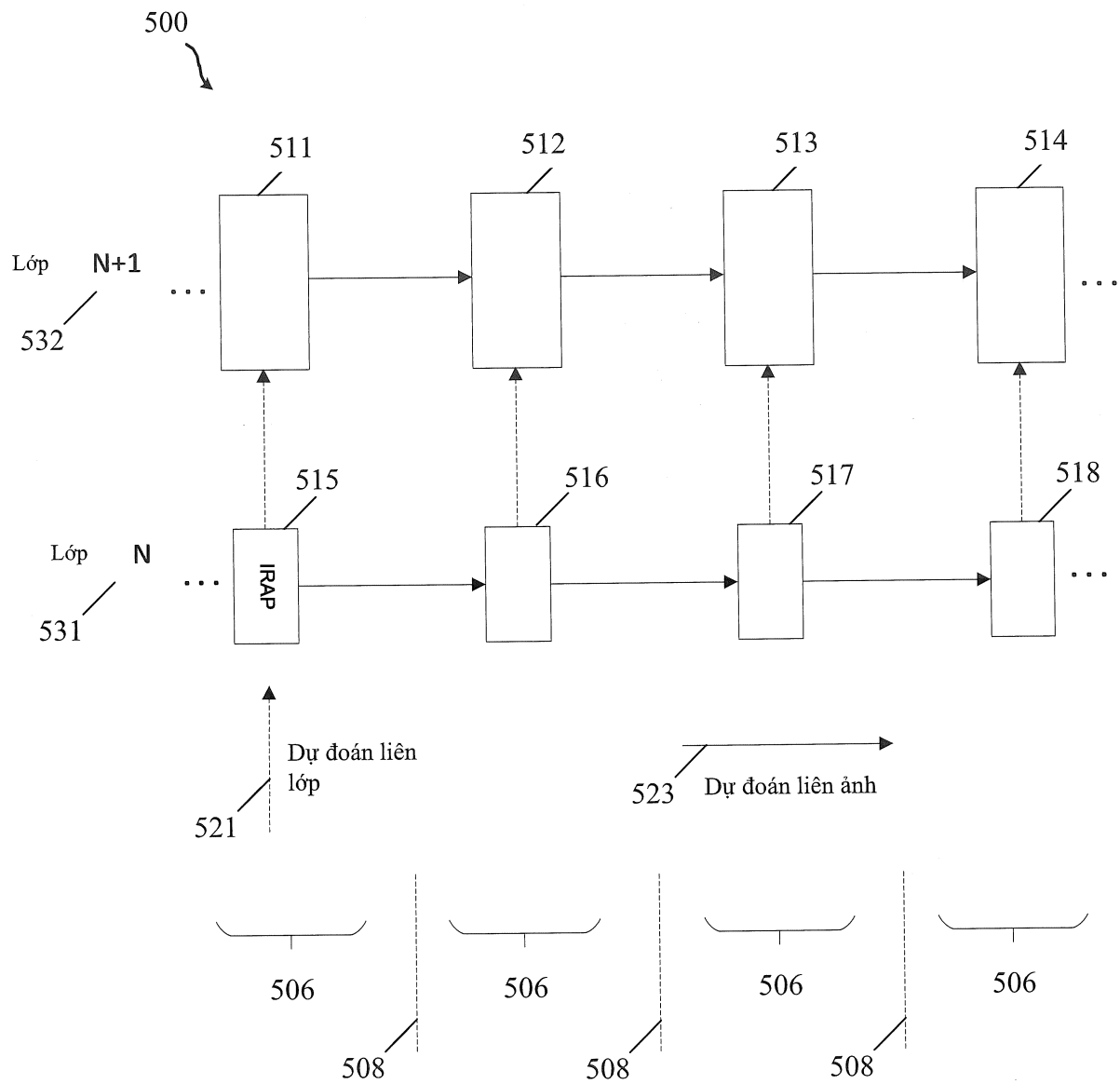


FIG. 5

6/13

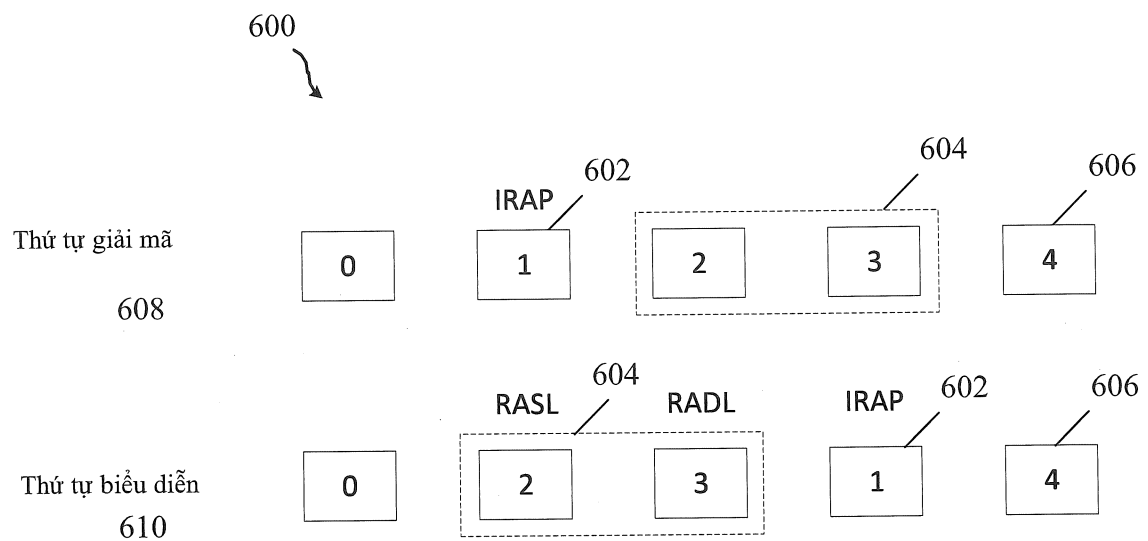


FIG. 6

7/13

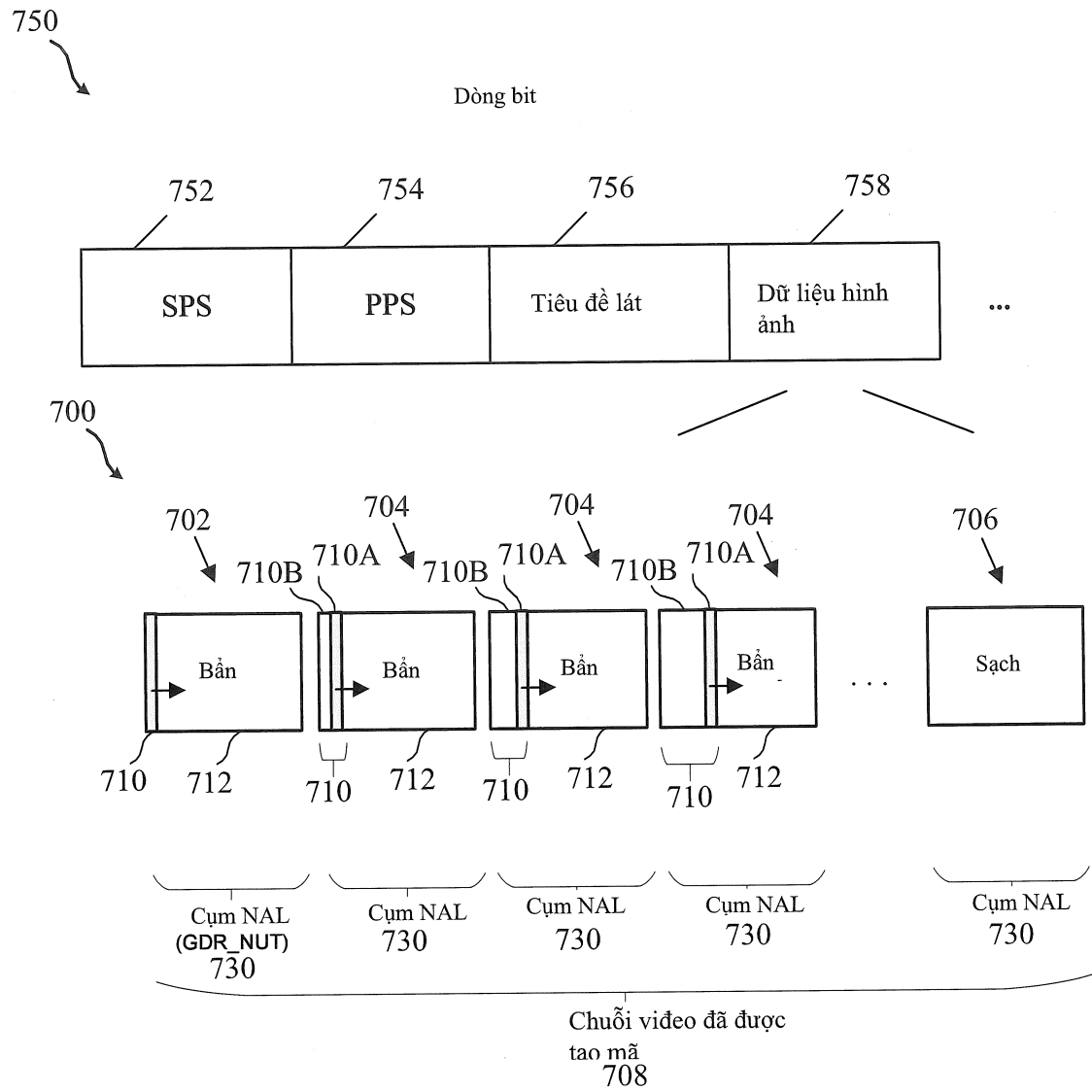


FIG. 7

8/13

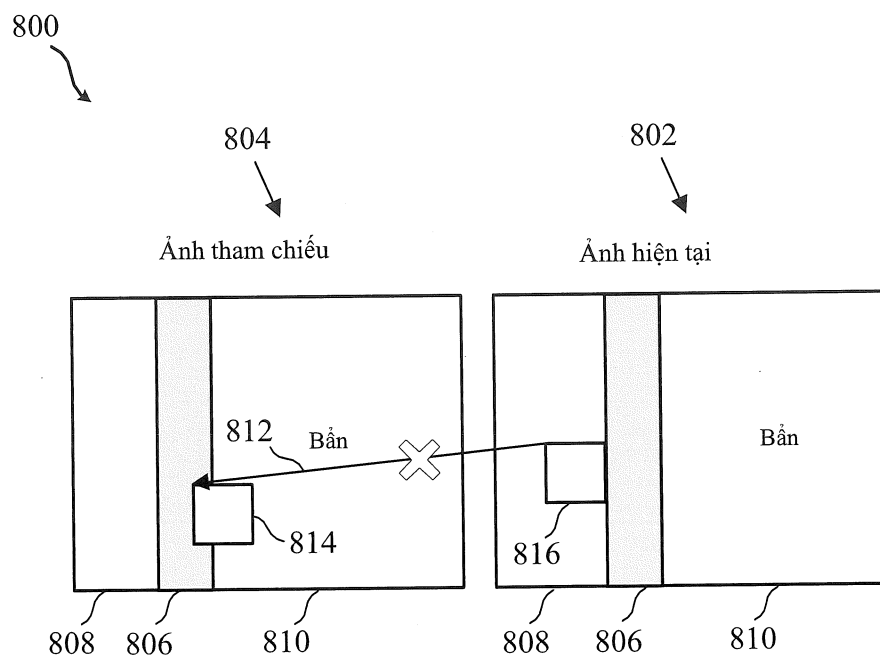


FIG. 8

9/13

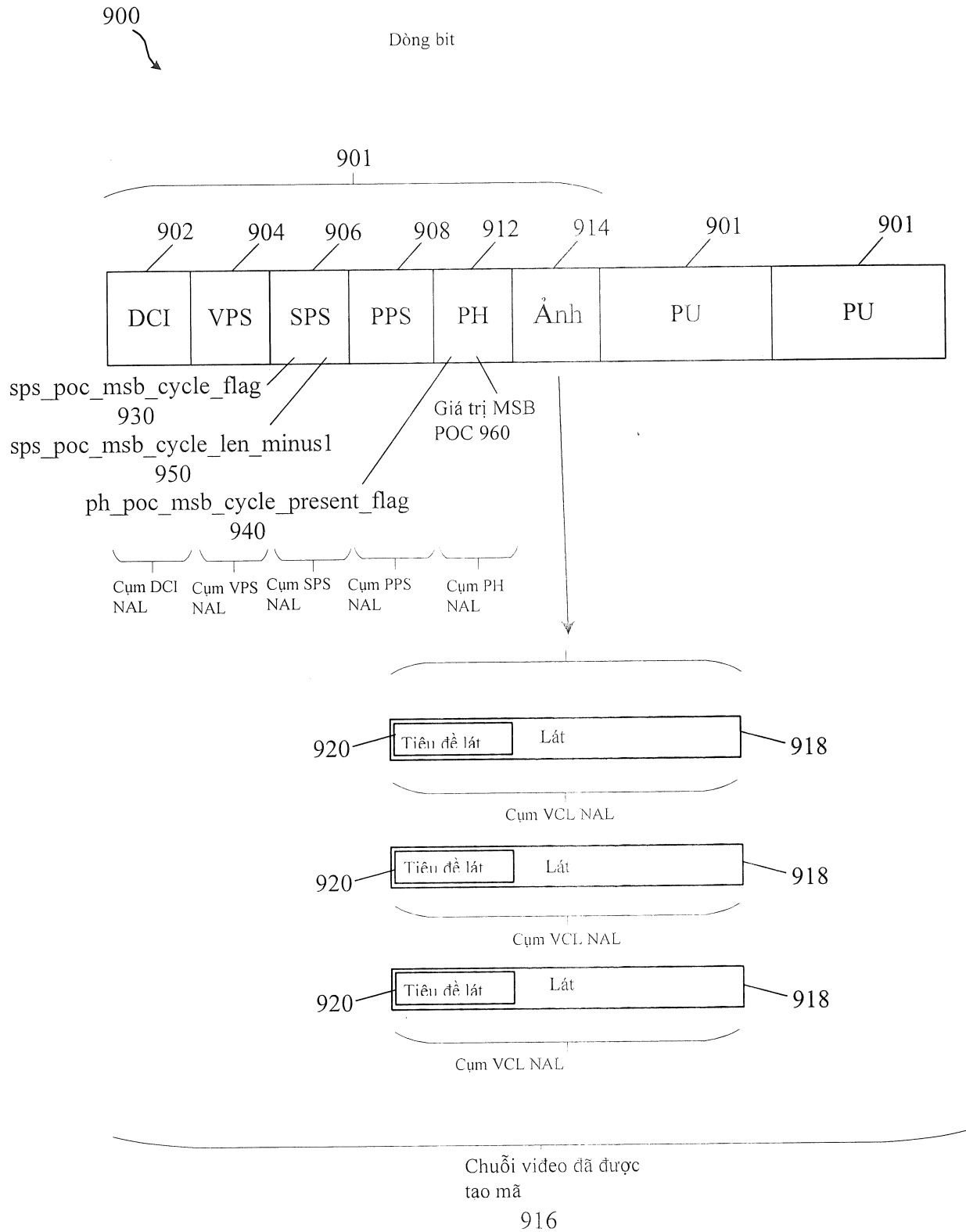


FIG. 9

10/13

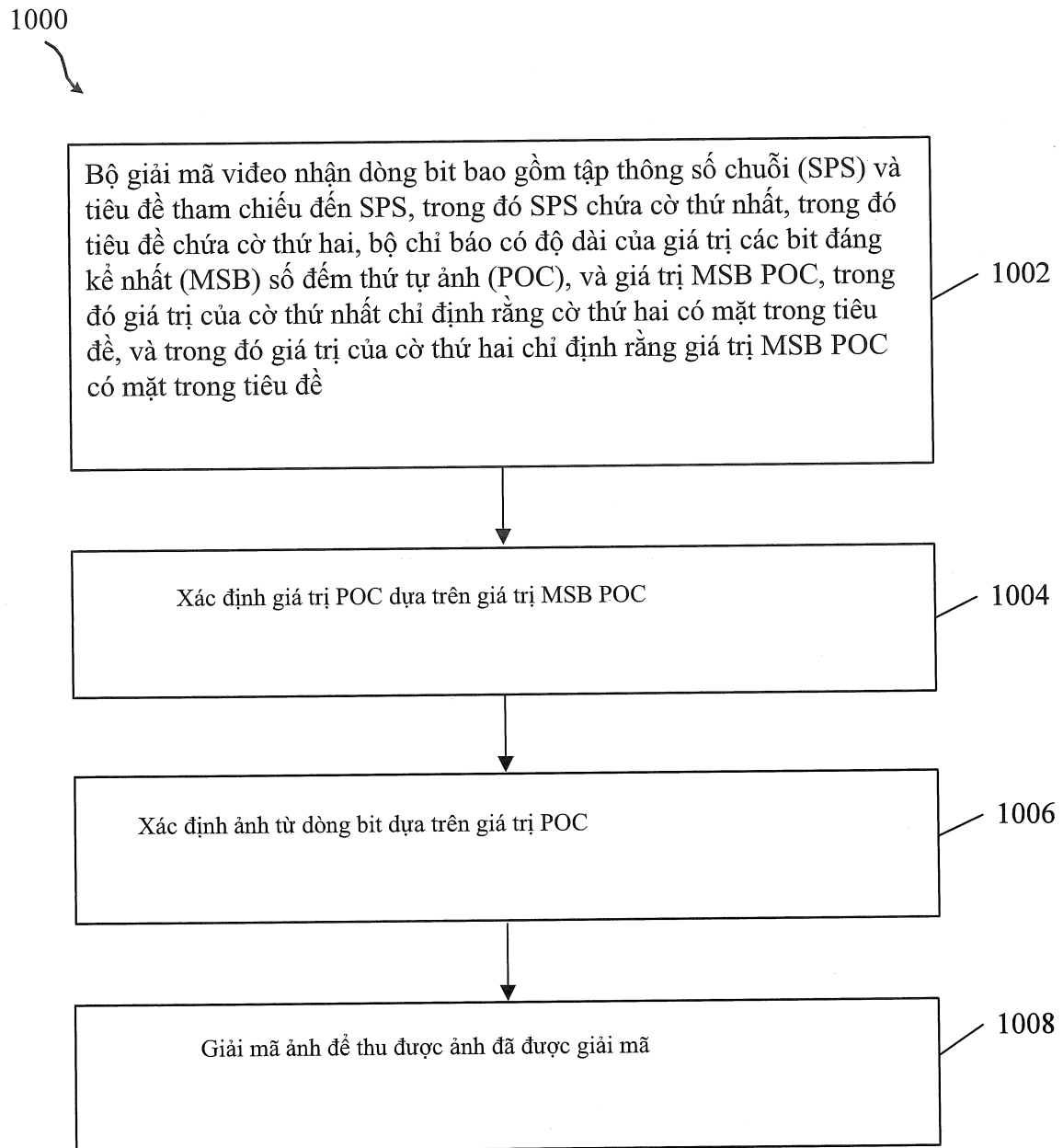


FIG. 10

11/13

1100

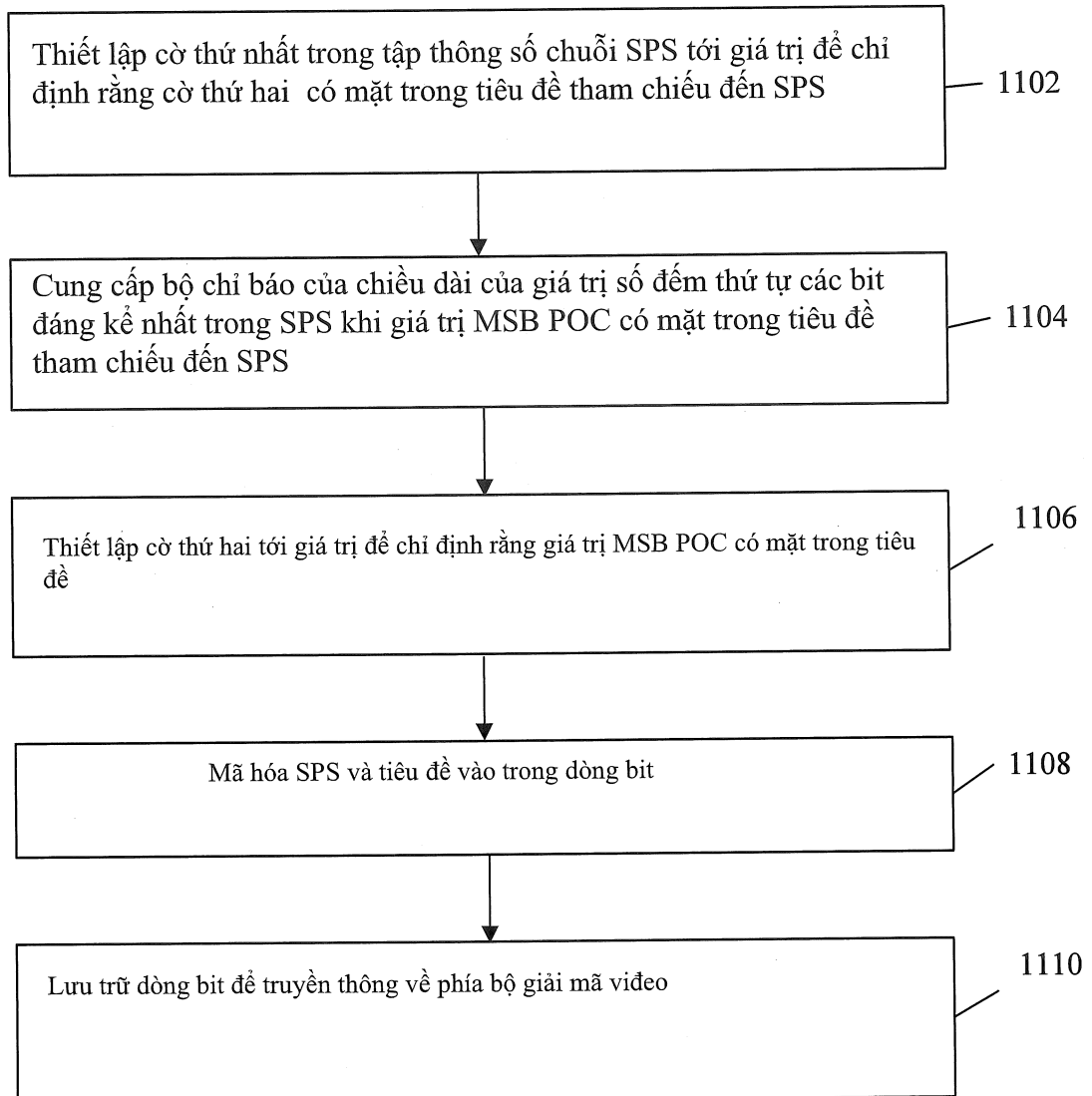


FIG. 11

12/13

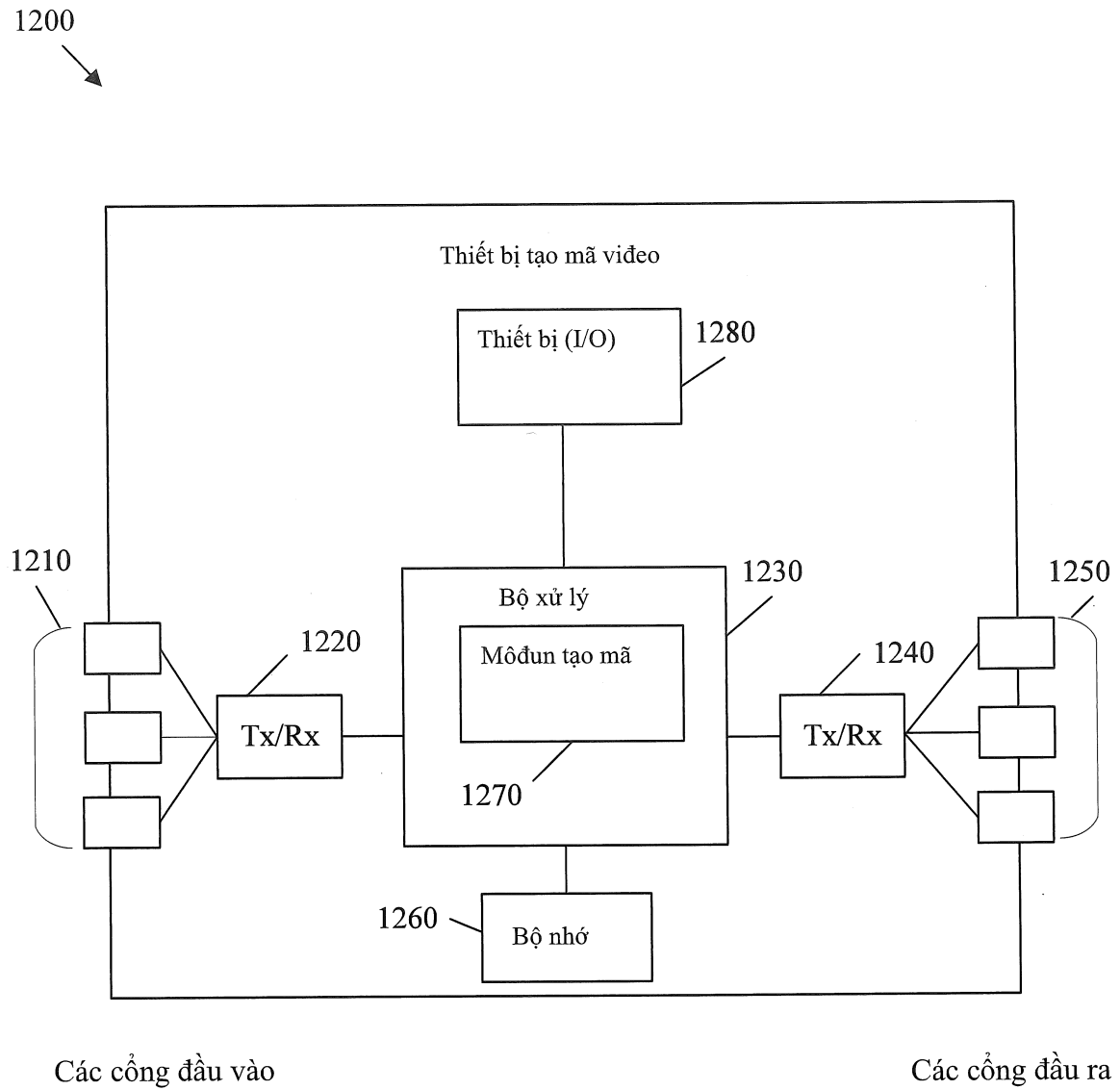


FIG. 12

13/13

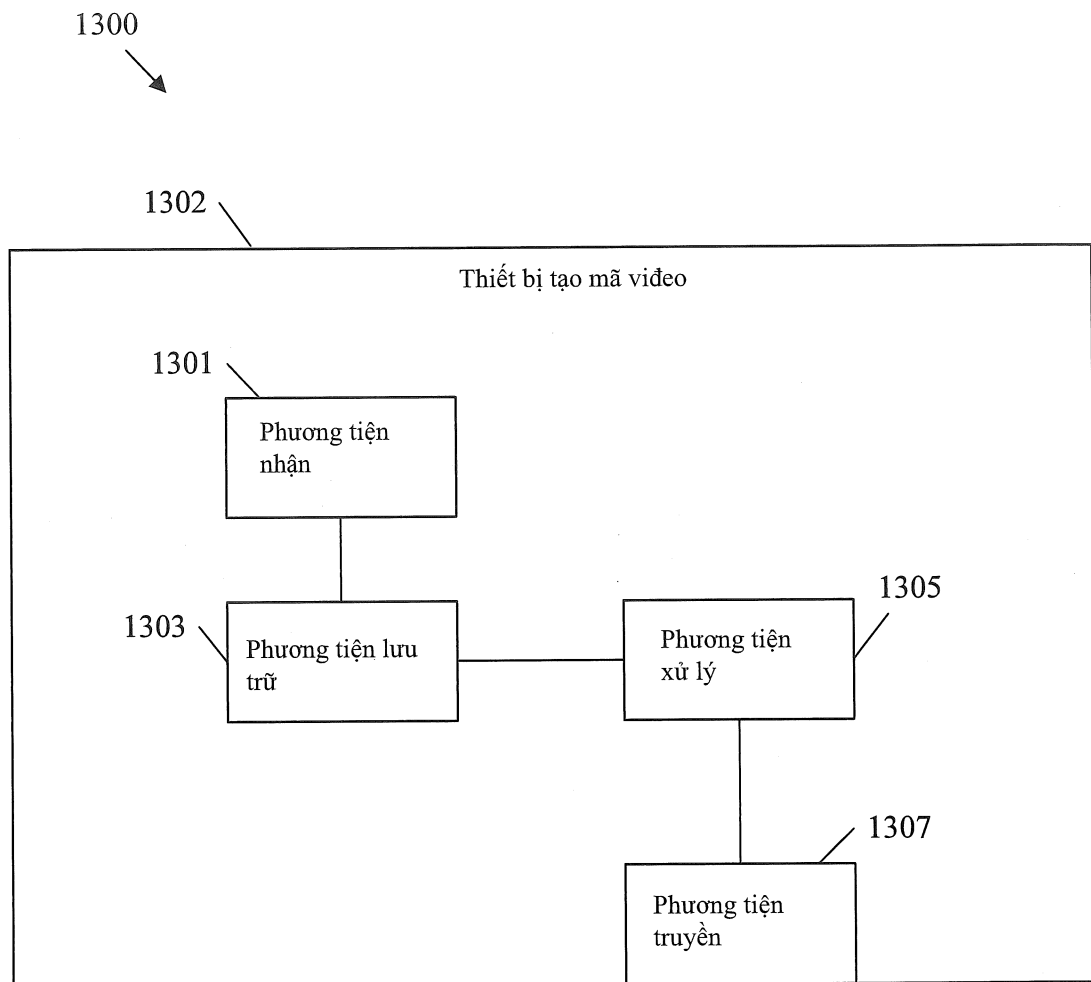


FIG. 13