



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

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11)
CỤC SỞ HỮU TRÍ TUỆ



1-0048116

(51)^{2020.01} H04N 19/159; H04N 19/172 (13) B

(21) 1-2021-06990

(22) 02/04/2020

(86) PCT/US2020/026457 02/04/2020

(87) WO2020/206171 08/10/2020

(30) 62/828,875 03/04/2019 US; 62/864,958 21/06/2019 US

(45) 25/07/2025 448

(43) 25/01/2022 406A

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

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

(72) HENDRY, Fnu (ID); WANG, Ye-Kui (US).

(74) Công ty Luật TNHH T&G (TGVN)

(54) PHƯƠNG PHÁP GIẢI MÃ, PHƯƠNG PHÁP MÃ HÓA, PHƯƠNG TIỆN ĐỌC
ĐƯỢC BẰNG MÁY TÍNH BẤT BIẾN, BỘ GIẢI MÃ, VÀ BỘ MÃ HÓA

(21) 1-2021-06990

(57) Sáng chế đề cập đến bộ mã hóa, bộ giải mã, thiết bị mã hóa video, và các phương pháp giải mã và tạo mã, phương tiện đọc được bằng máy tính bất biến và cơ chế tạo mã video. Cơ chế bao gồm bước nhận dòng bit bao gồm cờ và các ảnh được tạo mã bao gồm ảnh điểm truy nhập ngẫu nhiên trong ảnh (IRAP: intra random access point) và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP. Khi cờ được đặt thành giá trị thứ nhất, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự giải mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP. Khi cờ được đặt thành giá trị thứ hai, ảnh không dẫn ở trước, theo thứ tự giải mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP. Ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP được giải mã theo thứ tự giải mã dựa trên cờ.

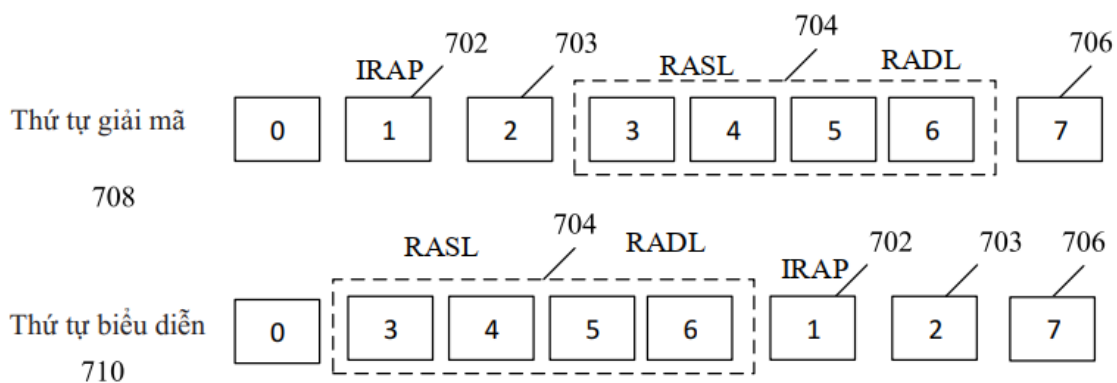


FIG. 7

Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế liên quan đến mã hóa video, và cụ thể là liên quan đến mã hóa các ảnh dẫn trong ngữ cảnh tạo mã video đan xen.

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

Theo một phương án thực hiện sáng chế, sáng chế bao gồm phương pháp được thực hiện trong bộ giải mã, phương pháp bao gồm các bước: nhận, bởi bộ nhận của bộ giải mã, dòng bit bao gồm cờ và các ảnh được tạo mã bao gồm ảnh điểm truy nhập ngẫu nhiên trong ảnh (IRAP: intra random access point) và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP; xác định, bởi bộ xử lý, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự giải mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP khi cờ được đặt thành giá trị thứ nhất; xác định, bởi bộ xử lý, ảnh không dẫn ở

trước, theo thứ tự giải mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP khi còn được đặt thành giá trị thứ hai; giải mã, bởi bộ xử lý, ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP theo thứ tự giải mã dựa trên việc còn được đặt thành giá trị thứ nhất hoặc giá trị thứ hai; và chuyển tiếp, bởi bộ xử lý, một hoặc nhiều ảnh đã được giải mã để hiển thị làm một phần của chuỗi video đã được giải mã.

Các hệ thống video tạo mã video vạn năng (VVC: Versatile Video Coding) có thể sử dụng dòng bit mà bao gồm các ảnh IRAP, các ảnh dẫn, và các ảnh không dẫn. Các ảnh không dẫn cũng có thể được gọi là các ảnh theo sau theo một số ví dụ. Ảnh IRAP là ảnh đã được tạo mã dự đoán trong ảnh mà có vai trò như sự bắt đầu của chuỗi video đã được tạo mã. Ảnh dẫn là ảnh mà ở trước ảnh IRAP theo thứ tự biểu diễn, nhưng được tạo mã sau ảnh IRAP theo thứ tự tạo mã. Ảnh không dẫn/ảnh theo sau là ảnh mà theo sau ảnh IRAP trong cả sự biểu diễn cũ hơn và hoạt động tạo mã cũ hơn. Một số hệ thống mã hóa video yêu cầu rằng các ảnh dẫn theo ngay sau ảnh IRAP theo thứ tự giải mã và rằng tất cả các ảnh không dẫn theo sau các ảnh dẫn. Hoạt động tạo mã video đan xen là cơ chế làm tăng tốc độ khung cảm nhận được mà không làm tăng băng thông phát trực tuyến. Trong hoạt động tạo mã video đan xen, khung video được chia thành hai trường. Các đường nằm ngang cho trường thứ nhất của khung được chụp ở thời điểm thứ nhất và được tạo mã trong ảnh thứ nhất. Các đường nằm ngang cho trường thứ hai của khung được chụp ở thời điểm thứ hai và được tạo mã trong ảnh thứ hai liền kề ngay với ảnh thứ nhất. Theo cách này, khung tạo thành bao gồm các lát cắt từ ảnh thứ nhất ở thời điểm thứ nhất và các lát cắt từ ảnh thứ hai ở thời điểm thứ hai, để làm tăng chiều của chuyển động. Các hệ thống VVC có thể không được thiết kế để hỗ trợ video đan xen. Ví dụ, khung đan xen có thể sử dụng ảnh IRAP và ảnh đã được tạo mã dự đoán trong ảnh liền kề để xử lý. Ảnh đã được tạo mã dự đoán trong ảnh được xem là các ảnh không dẫn/theo sau. Ngoài ra, khi các ảnh dẫn được sử dụng, các ảnh dẫn sẽ được định vị sau ảnh đã được tạo mã dự đoán trong ảnh liền kề đó. Điều này vi phạm ràng buộc VVC mà các ảnh dẫn theo ngay sau ảnh IRAP theo thứ tự giải mã và tất cả các ảnh không dẫn theo sau các ảnh dẫn. Ví dụ này bao gồm còi mà có thể được sử dụng để thực hiện hoạt động tạo mã video đan xen trong hệ thống VVC sử dụng các ảnh dẫn. Khi còn được đặt thành giá trị thứ nhất, như không, các ảnh dẫn, nếu có, ưu tiên tất cả các ảnh

không dẫn. Tuy nhiên, bộ mã hóa có thể đặt cờ thành giá trị thứ hai, như một, để chỉ báo cho bộ giải mã mà một ảnh không dẫn được định vị giữa ảnh IRAP và các ảnh dẫn bất kỳ. Theo một ví dụ, các ảnh không dẫn có thể không được định vị giữa các ảnh dẫn. Cờ có thể có mặt trong tập thông số chuỗi (SPS) và có thể áp dụng cho toàn bộ chuỗi của các ảnh. Theo đó, ví dụ này bao gồm cờ để làm tăng chức năng của bộ mã hóa và/hoặc bộ giải mã bằng cách cho phép các ảnh dẫn và video đan xen được thực hiện với nhau trong cùng dòng bit. Ngoài ra, ví dụ này làm tăng hiệu suất tạo mã của dòng bit tạo thành bằng cách cho phép các ảnh dẫn và video đan xen được thực hiện với nhau. Như vậy, ví dụ này có thể làm giảm việc sử dụng các tài nguyên bộ xử lý, bộ nhớ các tài nguyên, và/hoặc các tài nguyên mạng ở bộ mã hóa và/hoặc ở bộ giải mã.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, còn bao gồm bước xác định, bởi bộ xử lý, rằng không có các ảnh dẫn được định vị giữa ảnh dẫn ban đầu và ảnh dẫn cuối cùng theo thứ tự giải mã khi cờ được đặt thành giá trị thứ hai.

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, trong đó dòng bit bao gồm SPS, và trong đó cờ thu được từ SPS.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó cờ là cờ trường tuần tự (field_seq_flag).

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó field_seq_flag được đặt thành một khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường, và trong đó field_seq_flag được đặt thành không khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó ảnh IRAP bao gồm trường thứ nhất của khung, và trong đó ảnh không dẫn ở trước ảnh dẫn ban đầu bao gồm trường thứ hai của khung.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó bước giải mã ảnh IRAP và một hoặc

nhiều ảnh không dẫn bao gồm đan xen trường thứ nhất từ ảnh IRAP và trường thứ hai từ ảnh không dẫn ở trước ảnh dẫn ban đầu để tạo ra một khung.

Theo một phương án, sáng chế bao gồm phương pháp được thực hiện trong bộ mã hóa, phương pháp này bao gồm các bước: xác định, bởi bộ xử lý của bộ mã hóa, thứ tự tạo mã cho chuỗi video bao gồm các ảnh có ảnh IRAP và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP; mã hóa, bởi bộ xử lý, cò vào trong dòng bit, trong đó cò được đặt thành giá trị thứ nhất khi các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự tạo mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP, và trong đó cò được đặt thành giá trị thứ hai khi ảnh không dẫn ở trước, theo thứ tự tạo mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP; mã hóa, bởi bộ xử lý, ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP vào trong dòng bit theo thứ tự tạo mã; và lưu trữ, bởi bộ nhớ được ghép nối với bộ xử lý, dòng bit để truyền thông về phía bộ giải mã.

Các hệ thống video VVC có thể sử dụng dòng bit mà bao gồm các ảnh IRAP, các ảnh dẫn, và các ảnh không dẫn. Các ảnh không dẫn cũng có thể được gọi là các ảnh theo sau theo một số ví dụ. Ảnh IRAP là ảnh đã được tạo mã dự đoán trong ảnh mà có vai trò như sự bắt đầu của chuỗi video đã được tạo mã. Ảnh dẫn là ảnh mà ở trước ảnh IRAP theo thứ tự biểu diễn, nhưng được tạo mã sau ảnh IRAP theo thứ tự tạo mã. Ảnh không dẫn/ảnh theo sau là ảnh mà theo sau ảnh IRAP trong cả sự biểu diễn cũ hơn và hoạt động tạo mã cũ hơn. Một số hệ thống mã hóa video yêu cầu rằng các ảnh dẫn theo ngay sau ảnh IRAP theo thứ tự giải mã và rằng tất cả các ảnh không dẫn theo sau các ảnh dẫn. Hoạt động tạo mã video đan xen là cơ chế làm tăng tốc độ khung cảm nhận được mà không làm tăng băng thông phát trực tuyến. Trong hoạt động tạo mã video đan xen, khung video được chia thành hai trường. Các đường nằm ngang cho trường thứ nhất của khung được chụp ở thời điểm thứ nhất và được tạo mã trong ảnh thứ nhất. Các đường nằm ngang cho trường thứ hai của khung được chụp ở thời điểm thứ hai và được tạo mã trong ảnh thứ hai liền kề ngay với ảnh thứ nhất. Theo cách này, khung tạo thành bao gồm các lát cắt từ ảnh thứ nhất ở thời điểm thứ nhất và các lát cắt từ ảnh thứ hai ở thời điểm thứ hai, để làm tăng chiều của chuyển động. Các hệ thống VVC có thể không được thiết kế để hỗ trợ video đan xen. Ví dụ, khung đan xen có thể sử dụng ảnh IRAP và ảnh đã được tạo mã dự đoán trong ảnh liền kề để xử lý. Ảnh đã được tạo mã dự đoán trong

ảnh được xem là các ảnh không dẫn/theo sau. Ngoài ra, khi các ảnh dẫn được sử dụng, các ảnh dẫn sẽ được định vị sau ảnh đã được tạo mã dự đoán trong ảnh liền kề đó. Điều này vi phạm ràng buộc VVC mà các ảnh dẫn theo ngay sau ảnh IRAP theo thứ tự giải mã và tất cả các ảnh không dẫn theo sau các ảnh dẫn. Ví dụ này bao gồm cờ mà có thể được sử dụng để thực hiện hoạt động tạo mã video đan xen trong hệ thống VVC sử dụng các ảnh dẫn. Khi cờ được đặt thành giá trị thứ nhất, như không, các ảnh dẫn, nếu có, ưu tiên tất cả các ảnh không dẫn. Tuy nhiên, bộ mã hóa có thể đặt cờ thành giá trị thứ hai, như một, để chỉ báo cho bộ giải mã mà một ảnh không dẫn được định vị giữa ảnh IRAP và các ảnh dẫn bất kỳ. Theo một ví dụ, các ảnh không dẫn có thể không được định vị giữa các ảnh dẫn. Cờ có thể có mặt trong SPS và có thể áp dụng cho toàn bộ chuỗi của các ảnh. Theo đó, ví dụ này bao gồm cờ để làm tăng chức năng của bộ mã hóa và/hoặc bộ giải mã bằng cách cho phép các ảnh dẫn và video đan xen được thực hiện với nhau trong cùng dòng bit. Ngoài ra, ví dụ này làm tăng hiệu suất tạo mã của dòng bit tạo thành bằng cách cho phép các ảnh dẫn và video đan xen được thực hiện với nhau. Như vậy, ví dụ này có thể làm giảm việc sử dụng các tài nguyên bộ xử lý, bộ nhớ các tài nguyên, và/hoặc các tài nguyên mạng ở bộ mã hóa và/hoặc ở bộ giải mã.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh cung cấp, trong đó không có các ảnh dẫn được định vị giữa ảnh dẫn ban đầu và ảnh dẫn cuối cùng theo thứ tự tạo mã khi cờ được đặt thành giá trị thứ hai.

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, ở đây dòng bit bao gồm SPS, và trong đó cờ được mã hóa vào trong SPS.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó cờ là `field_seq_flag`.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó `field_seq_flag` được đặt thành một khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường, và trong đó `field_seq_flag` được đặt thành không khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó ảnh IRAP bao gồm trường thứ nhất của khung, và trong đó ảnh không dẫn ở trước ảnh dẫn ban đầu bao gồm trường thứ hai của khung.

Theo cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, cách thực hiện khác của khía cạnh này đề xuất, trong đó trường thứ nhất từ ảnh IRAP và trường thứ hai từ ảnh không dẫn ở trước ảnh dẫn ban đầu bao gồm việc xen kẽ các đường của dữ liệu video mà thể hiện một khung video đan xen.

Theo một phương án, sáng chế bao gồm thiết bị mã hóa video bao gồm: bộ xử lý, bộ nhận được ghép nối với bộ xử lý, bộ nhớ được ghép nối với bộ xử lý, và bộ truyền được ghép nối với bộ xử lý, trong đó bộ xử lý, bộ nhận, bộ nhớ, và bộ truyền được tạo cấu hình để thực hiện phương pháp theo khía cạnh bất kỳ trong số các khía cạnh nêu trên.

Theo một phương án thực hiện sáng chế, sáng chế bao gồm phương tiện đọc được bằng máy tính bất biến bao gồm sản phẩm chương trình máy tính để sử dụng bởi thiết bị mã hóa video, sản phẩm chương trình máy tính bao gồm các lệnh có thể thực thi bằng máy tính được lưu trữ trên phương tiện đọc được bằng máy tính bất biến sao cho khi được thực thi bởi bộ xử lý thì khiến cho thiết bị mã hóa video thực hiện phương pháp theo khía cạnh bất kỳ trong số các khía cạnh nêu trên.

Theo một phương án, sáng chế bao gồm bộ giải mã bao gồm: phương tiện nhận để nhận dòng bit bao gồm cờ và các ảnh được tạo mã bao gồm ảnh IRAP và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP; phương tiện xác định để: xác định các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự giải mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP khi cờ được đặt thành giá trị thứ nhất; và xác định ảnh không dẫn ở trước, theo thứ tự giải mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP khi cờ được đặt thành giá trị thứ hai; phương tiện giải mã để giải mã ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP theo thứ tự giải mã dựa trên việc cờ được đặt thành giá trị thứ nhất hoặc giá trị thứ hai; và phương tiện chuyển tiếp để chuyển tiếp một hoặc nhiều ảnh đã được giải mã để hiển thị làm một phần của chuỗi video đã được giải mã.

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, trong đó bộ giải mã còn được cấu hình để thực hiện phương pháp theo khía cạnh bất kỳ trong số các khía cạnh nêu trên.

Theo một phương án, sáng chế bao gồm bộ mã hóa bao gồm: phương tiện xác định để xác định thứ tự tạo mã cho chuỗi video bao gồm các ảnh có ảnh IRAP và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP; phương tiện mã hóa để: mã hóa cờ vào trong dòng bit, trong đó cờ được đặt thành giá trị thứ nhất khi các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự tạo mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP, và trong đó cờ được đặt thành giá trị thứ hai khi ảnh không dẫn ở trước, theo thứ tự tạo mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP; và mã hóa ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP vào trong dòng bit theo thứ tự tạo mã; và phương tiện lưu trữ để lưu trữ dòng bit để truyền thông về phía bộ giải mã.

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, trong đó bộ mã hóa còn được cấu hình để thực hiện phương pháp theo khía cạnh bất kỳ trong số các khía cạnh nêu trên.

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 là hình vẽ dạng giản lược thể hiện chuỗi video đã được tạo mã làm ví dụ với các ảnh dẫn.

FIG.6A đến Fig.6C là các hình vẽ dạng giản lược thể hiện chung một ví dụ về hoạt động tạo mã video đan xen.

FIG.7 là hình vẽ dạng giản lược thể hiện chuỗi video đã được tạo mã làm ví dụ sử dụng cả hoạt động tạo mã video đan xen và các ảnh dẫn.

FIG.8 là hình vẽ dạng giản lược thể hiện dòng bit làm ví dụ được tạo cấu hình để chứa cả hoạt động tạo mã video đan xen và các ảnh dẫn.

FIG.9 là hình vẽ dạng sơ đồ của thiết bị mã hóa video làm ví dụ.

FIG.10 là lưu đồ của phương pháp làm ví dụ để mã hóa chuỗi video với hoạt động tạo mã video đan xen và các ảnh dẫn vào trong dòng bit.

FIG.11 là lưu đồ của phương pháp làm ví dụ để giải mã chuỗi video với hoạt động tạo mã video đan xen và các ảnh dẫn từ dòng bit.

FIG.12 là hình vẽ dạng giản lược của hệ thống làm ví dụ để tạo mã chuỗi video với hoạt động tạo mã video đan xen và các ảnh dẫn vào trong dòng bit.

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 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. Cờ là bit hoặc nhóm của các bit được tạo mã vào trong dòng bit mà báo hiệu các cơ chế được sử dụng bởi bộ mã hóa trong mã hóa, và do đó chỉ báo các cơ chế được sử dụng bởi bộ giải mã trong quá trình giải mã một cách chính xác dữ liệu video tái cấu trúc từ dòng bit. Dự đoán trong ảnh là cơ chế tạo mã ảnh khi tham chiếu đến chính nó sao cho ảnh có thể được tái tạo mà không cần tham chiếu đến các ảnh khác. Dự đoán liên ảnh là cơ chế tạo mã ảnh nhờ tham chiếu đến một hoặc nhiều ảnh khác. Các ảnh điểm truy nhập ngẫu nhiên trong ảnh (IRAP) là ảnh mà được tạo mã theo dự đoán trong ảnh và có vai trò như điểm bắt đầu cho chuỗi video đã được tạo mã. Ảnh dẫn là ảnh được tạo mã sau khi ảnh IRAP kết hợp theo thứ tự tạo mã, nhưng ở trước ảnh IRAP kết hợp theo thứ tự xuất. Ảnh không dẫn, mà cũng có thể được gọi là ảnh theo sau, là ảnh mà theo sau ảnh IRAP theo cả thứ tự tạo mã và thứ tự xuất. Hoạt động tạo mã video đan xen là cơ chế tạo mã video để tạo mã trường thứ nhất của dữ liệu video ở thời điểm thứ nhất trong ảnh thứ nhất, tạo mã trường thứ hai của dữ liệu video ở thời điểm thứ hai trong ảnh thứ hai, và kết hợp trường thứ nhất và trường thứ hai thành một khung dùng cho việc biểu diễn để cung cấp ấn tượng về tốc độ khung tăng. Khung là hình ảnh hoàn thành mà được dự định cho việc hiển thị hoàn toàn hoặc một phần đối với người dùng ở thời điểm tương ứng trong chuỗi video. Ảnh là khung ngoại trừ trừ các ngữ cảnh liên quan đến video đan xen mà trong trường hợp đó ảnh là trường của khung. Tập thông số là phần của dòng bit mà báo hiệu dữ liệu, như các cờ và các thông số khác, cho đoạn tương ứng của chuỗi video đã được tạo mã. Cờ trường tuần tự (field_seq_flag) là cờ, mà được sử dụng cho video đan xen và các tín hiệu khi ảnh không dẫn được định vị giữa ảnh các ảnh IRAP và dẫn theo thứ tự tạo mã.

Các từ viết tắt sau đây được sử dụng ở đây, cây tạo mã khối (CTB: Coding Tree Block), cụm cây tạo mã (CTU: Coding Tree Unit), cụm tạo mã (CU: Coding Unit), chuỗi video đã được tạo mã (CVS: Coded Video Sequence), nhóm chuyên gia video

liên kết (JVET: Joint Video Experts Team), tập hợp ảnh xếp lớp được hạn chế chuyển động (MTCS: Motion-constrained tile set), cụm truyền tối đa (MTU: Maximum Transfer Unit), lớp trừu tượng mạng (NAL: network abstraction layer), số đếm thứ tự ảnh (POC: Picture Order Count), phần tải dữ liệu chuỗi byte thô (RBSP: Raw Byte Sequence Payload), tập thông số chuỗi (SPS), và bản thảo công tác (WD: Working Draft).

Nhiều công nghệ nén video có thể được sử dụng để làm giảm kích thước của các tệp video với mức thất thoát dữ liệu tối thiểu. Ví dụ, các công nghệ nén video có thể bao gồm thực hiện dự đoán không gian (ví dụ, trong ảnh) và/hoặc dự đoán thời gian (ví dụ, liên ảnh) để làm giảm hoặc loại bỏ dư thừa dữ liệu trong các chuỗi video. Đối với hoạt động tạo mã video dựa trên khối, lát cắt video (ví dụ, 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 tạo mã 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 dự đoán một chiều (P) hoặc dự đoán hai chiều (B) được tạo mã liên ảnh (P hoặc B) của ảnh có thể được tạo mã bằng cách 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/hoặc các hình ảnh và các ảnh tham chiếu có thể được gọi là các khung tham chiếu và/hoặc các hình ảnh tham chiếu. Dự đoán không gian hoặc thời gian tạo ra khối dự đoán đại diện cho khối hình ảnh. Dữ liệu dư thể hiện các khác nhau điểm ảnh giữa khối hình ảnh gốc và khối dự đoán. Theo đó, khối được tạo mã liên ảnh được mã hóa theo vectơ 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. Điều này dẫn tới các hệ số biến đổi dư, mà có thể được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa có thể được bố trí ban đầu trong mảng hai chiều. Các hệ số biến đổi được lượng tử hóa có thể được quét để tạo ra vectơ một chiều của các hệ số

biến đổi. Hoạt động tạo mã entropy có thể được áp dụng vào đạt được độ nén cao hơn nữa. Các công nghệ nén video này được thảo luận chi tiết hơn bên dưới.

Để đảm bảo video được mã hóa có thể được giải mã chính xác, video được mã hóa và được giải mã theo các tiêu chuẩn tạo mã video tương ứng. Các tiêu chuẩn tạo mã video bao gồm lĩnh vực chuẩn hóa hiệp hội viễn thông quốc tế (ITU: International Telecommunication Union) (ITU-T) H.261, tổ chức tiêu chuẩn hóa quốc tế/Ủy ban kỹ thuật điện tử quốc tế (ISO/IEC) nhóm chuyên gia ảnh động (MPEG: Motion Picture Experts Group)-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 ba chiều (3D: three dimensional) 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). Nhóm chuyên gia video liên kết (JVET) của ITU-T và ISO/IEC đã bắt đầu phát triển tiêu chuẩn mã hóa video được gọi là tạo mã video vạn năng (VVC: Versatile Video Coding). VVC được bao hàm trong bản thảo công tác (WD: Working Draft), vốn bao gồm JVET-M1001-v7.

Các hệ thống mã hóa video có thể mã hóa video bằng cách sử dụng các ảnh IRAP và các ảnh không IRAP. Các ảnh IRAP là các ảnh được tạo mã theo dự đoán trong ảnh mà có vai trò như các điểm truy nhập ngẫu nhiên cho chuỗi video. Ở dự đoán trong ảnh, các khối của ảnh được tạo mã nhờ tham chiếu đến các khối khác trong cùng một ảnh. Điều này là ngược lại với các ảnh không IRAP ở chỗ nó sử dụng dự đoán liên ảnh. Ở dự đoán liên ảnh, các khối của ảnh hiện tại được tạo mã nhờ tham chiếu đến các khối khác trong ảnh tham chiếu vốn khác với ảnh hiện tại. Do ảnh IRAP được tạo mã mà không tham chiếu đến các ảnh khác, ảnh IRAP có thể được giải mã mà trước tiên không cần giải mã các ảnh khác bất kỳ. Theo đó, bộ giải mã có thể bắt đầu giải mã chuỗi video ở ảnh IRAP bất kỳ. Ngược lại, ảnh không IRAP được tạo mã khi tham chiếu đến các ảnh khác, và do đó bộ giải mã nói chung không thể bắt đầu giải mã chuỗi video ở ảnh không IRAP. Các ảnh IRAP cũng có thể nạp mới bộ đệm hình ảnh đã được giải mã

(DPB: decoded picture buffer). Điều này do ảnh IRAP là điểm bắt đầu cho chuỗi video đã được tạo mã (CVS: coded video sequence), và các ảnh trong CVS không viện dẫn đến các ảnh trong CVS trước đó. Như vậy, các ảnh IRAP cũng có thể dừng các lỗi tạo mã liên quan đến dự đoán liên ảnh do các lỗi này không thể lan truyền thông qua ảnh IRAP. Tuy nhiên, các ảnh IRAP lớn hơn đáng kể các ảnh không IRAP từ quan điểm kích cỡ dữ liệu. Như vậy, chuỗi video nói chung bao gồm nhiều ảnh không IRAP với số lượng nhỏ hơn của các ảnh IRAP rải rác để cân bằng hiệu suất tạo mã với chức năng. Ví dụ, 60 CVS khung có thể bao gồm một ảnh IRAP và 59 ảnh không IRAP. Theo đó, các ảnh IRAP làm giảm hiệu quả nén trong dòng bit. Ngoài ra, sự có mặt của ảnh IRAP trong dòng bit gây ra sóng theo tốc độ bit. Điều gây bất lợi này cho hiệu quả nén gây ra một phần do thực tế là dự đoán trong ảnh sử dụng nhiều bit hơn đáng kể để biểu diễn ảnh so với dự đoán liên ảnh. Ngoài ra, các ảnh IRAP có thể nạp mới quy trình giải mã và loại bỏ các ảnh tham chiếu từ DPB. Điều này làm giảm số lượng các ảnh tham chiếu khả dụng cho dự đoán liên ảnh khi tạo mã các ảnh theo sau ảnh IRAP, và do đó làm giảm tạm thời hiệu quả của quy trình dự đoán liên ảnh.

Các hệ thống mã hóa video cũng có thể sử dụng các ảnh dẫn. Ảnh dẫn là ảnh mà được định vị sau ảnh IRAP theo thứ tự tạo mã và trước ảnh IRAP theo thứ tự biểu diễn. Các ảnh dẫn có thể được sử dụng khi ảnh tương ứng có thể được dự đoán một cách hiệu quả từ ảnh IRAP mặc dù ảnh này nên được thể hiện trước ảnh IRAP. Ảnh này được định vị sau ảnh IRAP theo thứ tự tạo mã để cho phép ảnh IRAP được sử dụng làm ảnh tham chiếu cho dự đoán liên ảnh. Sau đó, bộ giải mã có thể hoán đổi thứ tự của ảnh dẫn và ảnh IRAP trước khi biểu diễn để tạo ra thứ tự biểu diễn khác nhau. Các ảnh dẫn có thể bao gồm các ảnh dẫn được bỏ qua truy nhập ngẫu nhiên (RASL) và các ảnh dẫn có thể giải mã truy nhập ngẫu nhiên (RADL). Ảnh RASL cũng có thể dựa vào các ảnh ở trước ảnh IRAP và được bỏ qua khi ảnh IRAP được sử dụng như điểm truy nhập ngẫu nhiên. Điều này là do các ảnh khác đó không được giải mã, và do đó không khả dụng như các ảnh tham chiếu khác khi việc giải mã bắt đầu từ ảnh IRAP. Ảnh RADL chỉ dựa vào ảnh IRAP hoặc các ảnh khác giữa ảnh RADL và ảnh IRAP để tham khảo. Do đó, các ảnh RADL được giải mã ngay cả khi IRAP được sử dụng như điểm truy nhập ngẫu nhiên. Điều này là do ảnh bất kỳ mà ảnh RADL có thể tham chiếu đến được bảo đảm sẽ được giải mã ngay cả khi việc tạo mã bắt đầu ở ảnh IRAP. Các hệ thống mã hóa video

có thể đòi hỏi rằng các ảnh dẫn được định vị ngay sau ảnh IRAP mà chúng tham chiếu đến theo thứ tự giải mã. Sau đó các ảnh theo sau được kết hợp bất kỳ theo sau các ảnh dẫn theo thứ tự giải mã.

Hoạt động tạo mã video sử dụng các cơ chế trong phạm vi rộng. Ví dụ, hoạt động tạo mã đan xen tạo mã khung vào trong nhiều hơn một trường và vào trong nhiều hơn một ảnh. Ví dụ, khung có thể được chia thành trường chẵn và trường lẻ. Trường chẵn của khung đan xen chứa các mẫu từ các đường nằm ngang được đánh số chẵn của khung và trường lẻ của khung đan xen chứa các mẫu từ các đường nằm ngang được đánh số lẻ của khung. Theo ví dụ cụ thể, trường chẵn có thể được chụp ở thời điểm thứ nhất và được lưu trữ trong ảnh thứ nhất. Sau đó, trường lẻ có thể được chụp ở thời điểm thứ hai và được lưu trữ trong ảnh thứ hai. Việc có hai trường trong cùng một khung làm tăng cảm giác chuyển động. Như vậy, việc tạo mã đan xen tạo ra ấn tượng về tốc độ khung tăng mà không làm tăng băng thông của chuỗi video. Việc tạo mã đan xen có thể không được hỗ trợ một cách tự nhiên bởi các hệ thống mã hóa được chuẩn hóa. Tuy nhiên, việc tạo mã đan xen có thể được quản lý trong một số hệ thống bằng cách sử dụng các phần tử cú pháp trong thông tin về khả năng sử dụng video (VUI) để chỉ báo dòng bit là dòng bit được tạo mã đan xen. Các phần tử cú pháp đó có thể bao gồm `field_seq_flag` và `general_frame_only_constraint_flag`.

Các hệ thống mã hóa video được chuẩn hóa mà sử dụng các ảnh dẫn không được tạo cấu hình để hỗ trợ hoạt động tạo mã video đan xen. Ví dụ, VVC và HEVC có thể sử dụng thứ tự tạo mã mà yêu cầu ảnh IRAP được theo sau bởi các ảnh dẫn, nếu có. Sau đó các ảnh dẫn theo sau bởi các ảnh không dẫn/theo sau. Thứ tự này ngăn không cho các ảnh không dẫn bị định vị giữa ảnh IRAP và các ảnh dẫn kết hợp. Tuy nhiên, trong ngữ cảnh tạo mã video đan xen, khung IRAP được chia giữa hai trường trong hai ảnh. Ảnh thứ nhất với trường thứ nhất được tạo mã như ảnh IRAP. Ảnh thứ hai với trường thứ hai được tạo mã như ảnh không dẫn/theo sau thay vì là ảnh IRAP do ảnh thứ hai không thể được sử dụng làm điểm truy nhập ngẫu nhiên. Điều này là do cả hai ảnh cần bắt đầu giải mã, và do đó ảnh thứ nhất không thể được bỏ qua. Hai ảnh mà tạo nên khung IRAP nên được định vị kế tiếp với nhau để tạo mã hiệu quả. Tuy nhiên, việc định vị ảnh không dẫn với trường IRAP thứ hai liền kề với ảnh IRAP với trường IRAP thứ

nhất vi phạm thứ tự tạo mã VVC và HEVC. Điều này là do việc định vị các vị trí ảnh không dẫn trước các ảnh dẫn bất kỳ.

Được bộc lộ ở đây là các cơ chế để tạo cấu hình hệ thống tạo mã video mà sử dụng các ảnh dẫn để mã hóa video đan xen. Ví dụ, cờ có thể được sử dụng để thực hiện hoạt động tạo mã video đan xen vào trong hệ thống VVC có sử dụng các ảnh dẫn. Cờ có thể được sử dụng để báo hiệu cho bộ giải mã khi ảnh không dẫn có thể có mặt giữa ảnh IRAP và các ảnh dẫn bất kỳ. Bộ giải mã có thể đọc cờ và điều chỉnh thứ tự theo mong muốn để hỗ trợ hoạt động tạo mã video đan xen. Khi cờ được đặt thành giá trị thứ nhất, như không, các ảnh dẫn, nếu có, ưu tiên tất cả các ảnh không dẫn. Tuy nhiên, bộ mã hóa có thể đặt cờ thành giá trị thứ hai, như một, để chỉ báo cho bộ giải mã mà một ảnh không dẫn được định vị giữa ảnh IRAP và các ảnh dẫn bất kỳ. Theo một ví dụ, các ảnh không dẫn có thể không được định vị giữa các ảnh dẫn. Ví dụ, cờ trường tuần tự (`field_seq_flag`) có thể được sử dụng nhằm đạt mục đích này. Cờ có thể có mặt trong tập thông số chuỗi (SPS) và có thể áp dụng cho toàn bộ chuỗi của các ảnh. Cần lưu ý rằng, trong ngữ cảnh video đan xen, khung có thể bao gồm nhiều ảnh (ví dụ, hai). Tuy nhiên, ngoài ngữ cảnh video đan xen, khung bao gồm một ảnh và do đó thuật ngữ khung và thuật ngữ ảnh có thể được sử dụng hoán đổi cho nhau. Theo đó, việc sử dụng của các thuật ngữ khung và ảnh bên dưới không nên được coi là làm giới hạn trừ khi được sử dụng trong ngữ cảnh tạo mã đan xen.

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 trị số 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 liền kề. 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 liền kề 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 a 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 nhiễu, 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

chiều 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 phần 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 trị số 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 có thể được phép (ví dụ, một bin cho hai tùy chọn, hai bin cho ba đến bốn tùy chọn, v.v.) Bộ mã hóa sau đó 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 vectơ 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 vectơ chuyển động, v.v. cho khối hiện tại/khung, và có thể chọn các khối tham chiếu, các vectơ 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 trị số 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 trị số 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 vectơ chuyển động với độ chính xác điểm ảnh phân số. Thành phần ước tính chuyển động 221 tính toán vectơ 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 vectơ 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 vectơ 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 vectơ chuyển động cho PU của khối video hiện tại, thành phần bù chuyển động 219 có thể định vị khối dự đoán mà vectơ chuyển động hướng vào đó. Sau đó khối video dư được tạo bằng cách trừ đi các trị số điểm ảnh của khối dự

đoán từ các trị số điểm ảnh của khối video hiện tại được mã hóa, tạo ra các trị số 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 trị số 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 trị số 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 a 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 trị số 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 vectơ 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.

FIG.5 là hình vẽ dạng giản lược thể hiện CVS 500 làm ví dụ với các ảnh dẫn. Ví dụ, CVS 500 có thể được mã hóa bởi bộ mã hóa, như hệ thống mã hóa 200 và/hoặc bộ mã hóa 300, theo phương pháp 100. Ngoài ra, CVS 500 có thể được giải mã bởi bộ giải mã, như hệ thống mã hóa 200 và/hoặc bộ giải mã 400. CVS 500 bao gồm các ảnh được tạo mã theo thứ tự giải mã 508. Thứ tự giải mã 508 là thứ tự trong đó các ảnh được định vị trong dòng bit. Sau đó các ảnh của CVS 500 được xuất ra theo thứ tự biểu diễn 510. Thứ tự biểu diễn 510 là thứ tự trong đó các ảnh nên được hiển thị bởi bộ giải mã để khiến cho video tạo thành hiển thị một cách thích hợp. Ví dụ, các ảnh của CVS 500 nói chung có thể được định vị theo thứ tự biểu diễn 510. Tuy nhiên, các ảnh nhất định có thể được dịch chuyển vào các vị trí khác nhau để tăng hiệu suất tạo mã, ví dụ bằng cách đặt các ảnh tương tự ở lân cận gần hơn để hỗ trợ việc dự đoán liên ảnh. Việc dịch chuyển các ảnh đó theo cách này dẫn tới thứ tự giải mã 508. Theo ví dụ được thể hiện, các ảnh được tạo chỉ mục theo thứ tự giải mã 508 từ không đến bốn. Theo thứ tự biểu diễn 510, các ảnh có chỉ số hai và chỉ số ba được dịch chuyển về phía trước ảnh có chỉ số không.

CVS 500 bao gồm ảnh IRAP 502. Ảnh IRAP 502 là ảnh được tạo mã theo dự đoán trong ảnh mà có vai trò làm điểm truy nhập ngẫu nhiên cho CVS 500. Cụ thể là, các khối của ảnh IRAP 502 được tạo mã nhờ tham chiếu đến các khối khác của ảnh IRAP 502. Do ảnh IRAP 502 được tạo mã mà không tham chiếu đến các ảnh khác, ảnh IRAP 502 có thể được giải mã mà trước tiên không cần giải mã các ảnh khác bất kỳ. Theo đó, bộ giải mã có thể bắt đầu giải mã CVS 500 ở ảnh IRAP 502. Ngoài ra, ảnh IRAP 502 có thể khiến cho DPB được nạp lại. Ví dụ, hầu hết các ảnh được thể hiện sau ảnh IRAP 502 có thể không dựa vào các ảnh ở trước ảnh IRAP 502 (ví dụ, chỉ số ảnh không) cho dự đoán liên ảnh. Như vậy, bộ đệm ảnh có thể được nạp lại ngay khi ảnh IRAP 502 được giải mã. Điều này có tác dụng dừng các lỗi tạo mã liên quan đến dự đoán liên ảnh bất kỳ do các lỗi này không thể lan truyền thông qua ảnh IRAP 502. Ảnh IRAP 502 có thể bao gồm các kiểu ảnh khác nhau. Ví dụ, ảnh IRAP có thể được tạo mã làm ảnh nạp lại bộ giải mã tức thời (IDR) hoặc truy nhập ngẫu nhiên sạch (CRA). IDR là ảnh được tạo mã trong ảnh mà bắt đầu CVS 500 mới và nạp mới bộ đệm ảnh. CRA là ảnh được tạo mã trong ảnh mà có vai trò như điểm truy nhập ngẫu nhiên mà không

cần bắt đầu CVS 500 mới hoặc nạp mới bộ đệm ảnh. Theo cách này, các ảnh dẫn 504 được liên kết với CRA có thể tham chiếu các ảnh trước CRA, trong khi các ảnh dẫn 504 được liên kết với IDR có thể không tham chiếu các ảnh trước IDR.

CVS 500 cũng bao gồm các ảnh không IRAP khác nhau. Các ảnh này bao gồm các ảnh dẫn 504 và các ảnh theo sau 506. Ảnh dẫn 504 là ảnh được định vị sau ảnh IRAP 502 theo thứ tự giải mã 508, nhưng được định vị trước ảnh IRAP 502 theo thứ tự biểu diễn 510. Các ảnh theo sau 506 được định vị sau ảnh IRAP 502 theo cả thứ tự giải mã 508 và thứ tự biểu diễn 510. Cả các ảnh dẫn 504 và các ảnh theo sau 506 cùng được tạo mã theo dự đoán liên ảnh trong hầu hết các trường hợp. Các ảnh theo sau 506 được tạo mã khi tham chiếu đến ảnh IRAP 502 hoặc các ảnh được định vị sau ảnh IRAP 502. Do đó, các ảnh theo sau 506 có thể luôn được giải mã ngay khi ảnh IRAP 502 được giải mã. Các ảnh dẫn 504 có thể bao gồm các ảnh dẫn được bỏ qua truy nhập ngẫu nhiên (RASL) và các ảnh dẫn có thể giải mã truy nhập ngẫu nhiên (RADL). Ảnh RASL được tạo mã nhờ tham chiếu đến các ảnh ở trước ảnh IRAP 502, nhưng được tạo mã ở vị trí sau ảnh IRAP 502. Do các ảnh RASL dựa vào các ảnh trước đó, ảnh RASL không thể được giải mã khi bộ giải mã bắt đầu việc giải mã ở ảnh IRAP 502. Theo đó, các ảnh RASL được bỏ qua và không được giải mã khi ảnh IRAP 502 được sử dụng như điểm truy nhập ngẫu nhiên. Tuy nhiên, các ảnh RASL được giải mã và được hiển thị khi bộ giải mã sử dụng ảnh IRAP trước đó (trước chỉ số không và không được thể hiện trên hình vẽ) làm điểm truy nhập ngẫu nhiên. Các ảnh RADL được tạo mã khi tham chiếu đến ảnh IRAP 502 và/hoặc các ảnh theo sau ảnh IRAP 502, nhưng được định vị trước ảnh IRAP 502 theo thứ tự biểu diễn. Do các ảnh RADL không dựa vào các ảnh ở trước ảnh IRAP 502, các ảnh RADL có thể được giải mã và được hiển thị khi ảnh IRAP 502 là điểm truy nhập ngẫu nhiên.

FIG.6A đến Fig.6C là các hình vẽ dạng giản lược thể hiện chung một ví dụ về hoạt động tạo mã video đan xen. Hoạt động tạo mã video đan xen tạo ra khung video đan xen 600, như được thể hiện trên Fig. 6C, từ ảnh thứ nhất 601 và ảnh thứ hai 602, như được thể hiện trên FIG.6A và FIG.6B. Ví dụ, hoạt động tạo mã video đan xen có thể được sử dụng bởi bộ mã hóa, như hệ thống mã hóa 200 và/hoặc bộ mã hóa 300, khi mã hóa video chứa khung video đan xen 600 làm một phần của phương pháp 100. Ngoài ra, bộ giải mã, như hệ thống mã hóa 200 và/hoặc bộ giải mã 400, có thể giải mã video

bao gồm khung video đan xen 600. Ngoài ra, khung video đan xen 600 có thể được mã hóa vào trong CVS, như CVS 500, như được mô tả chi tiết hơn đối với FIG.7 bên dưới.

Khi thực hiện hoạt động tạo mã video đan xen, trường thứ nhất 610 được chụp ở thời điểm thứ nhất và được mã hóa vào trong ảnh thứ nhất 601 như được thể hiện trên Fig.. 6A. Trường thứ nhất 610 bao gồm các đường nằm ngang của dữ liệu video. Cụ thể là, các đường nằm ngang của dữ liệu video trong trường thứ nhất 610 kéo dài từ đường bao bên trái của ảnh thứ nhất 601 sang đường bao bên phải của ảnh thứ nhất 601. Tuy nhiên, trường thứ nhất 610 bỏ qua các hàng xen kẽ của dữ liệu video. Theo một cách thực hiện làm ví dụ, trường thứ nhất 610 chứa một nửa dữ liệu video được chụp bởi thiết bị quay video ở thời điểm thứ nhất.

Như được thể hiện trên FIG.6B, trường thứ hai 612 được chụp ở thời điểm thứ hai và được mã hóa vào trong ảnh thứ hai 602. Ví dụ, thời điểm thứ hai có thể theo ngay sau thời điểm thứ nhất bởi giá trị được đặt dựa trên tốc độ khung được đặt cho video. Ví dụ, trong video được đặt để hiển thị ở khung tốc độ của mười lăm khung trên giây (FPS), thời điểm thứ hai có thể xảy ra một phần mười lăm giây sau thời điểm thứ nhất. Như được thể hiện, trường thứ hai 612 bao gồm các đường nằm ngang của dữ liệu video mà được bổ sung cho các đường nằm ngang của trường thứ nhất 610 của ảnh thứ nhất 601. Cụ thể là, các đường nằm ngang của dữ liệu video trong trường thứ hai 612 kéo dài từ đường bao bên trái của ảnh thứ hai 602 sang đường bao bên phải của ảnh thứ hai 602. Trường thứ hai 612 chứa các đường nằm ngang được bỏ qua bởi trường thứ nhất 610. Ngoài ra, trường thứ hai 612 bỏ qua các đường nằm ngang chứa trong trường thứ nhất 610.

Trường thứ nhất 610 của ảnh thứ nhất 601 và trường thứ hai 612 của ảnh thứ hai 602 có thể được kết hợp để hiển thị ở bộ giải mã như khung video đan xen 600 như được thể hiện trên FIG.6C. Cụ thể là, khung video đan xen 600 chứa trường thứ nhất 610 của ảnh thứ nhất 601 được chụp ở thời điểm thứ nhất và trường thứ hai 612 của ảnh thứ hai 602 được chụp ở thời điểm thứ hai. Sự kết hợp này có hiệu ứng hình ảnh để nhấn mạnh và/hoặc cường điệu sự chuyển động. Khi được hiển thị như một phần của video, chuỗi các khung video đan xen 600 tạo ra ấn tượng rằng video được mã hóa ở tốc độ khung tăng mà không cần mã hóa một cách thực tế các khung bổ sung. Theo cách này, hoạt động tạo mã video đan xen mà sử dụng khung video đan xen 600 có thể làm tăng

tốc độ khung hiệu dụng của video mà không kèm theo sự gia tăng về kích cỡ dữ liệu video. Như vậy, hoạt động tạo mã video đan xen có thể làm tăng hiệu suất tạo mã của chuỗi video được mã hóa.

FIG.7 là hình vẽ dạng giản lược thể hiện ví dụ CVS 700 sử dụng cả hoạt động tạo mã video đan xen, ví dụ để tạo ra khung video đan xen 600, và các ảnh dẫn. CVS 700 hầu như tương tự với CVS 500, nhưng được thay đổi để mã hóa các ảnh với các trường, như ảnh thứ nhất 601 và ảnh thứ hai 602, trong khi giữ lại các ảnh dẫn. Ví dụ, CVS 700 có thể được mã hóa bởi bộ mã hóa, như hệ thống mã hóa 200 và/hoặc bộ mã hóa 300, theo phương pháp 100. Ngoài ra, CVS 700 có thể được giải mã bởi bộ giải mã, như hệ thống mã hóa 200 và/hoặc bộ giải mã 400.

CVS 700 có thứ tự giải mã 708 và thứ tự biểu diễn 710, mà hoạt động theo cách hầu như tương tự với thứ tự giải mã 508 và thứ tự biểu diễn 510, một cách tương ứng. CVS 700 cũng chứa ảnh IRAP 702, các ảnh dẫn 704, và các ảnh theo sau 706, mà tương tự với ảnh IRAP 502, các ảnh dẫn 504, và các ảnh theo sau 506. Sự khác biệt là tất cả ảnh IRAP 702, các ảnh dẫn 704, và các ảnh theo sau 706 được tạo mã bằng cách sử dụng các trường theo cách hầu như tương tự với trường thứ nhất 610 và trường thứ hai 612 như đã được mô tả đối với FIG.6A đến Fig.6C. Như vậy, mỗi khung bao gồm hai ảnh. Theo đó, CVS 700 chứa ảnh nhiều hơn hai lần so với CVS 500. Tuy nhiên, CVS 700 chứa lượng dữ liệu gần như bằng với CVS 500 do mỗi trong số các ảnh của CVS 700 bỏ qua một nửa khung.

Vấn đề với CVS 700 là ảnh IRAP 702 được mã hóa bằng cách chứa trường thứ nhất của dự đoán trong ảnh được tạo mã dữ liệu. Sau đó trường thứ hai của dữ liệu được tạo mã dự đoán trong ảnh được chứa trong ảnh không dẫn 703. Ảnh không dẫn 703 không là ảnh IRAP 702 do bộ giải mã không thể bắt đầu giải mã CVS 700 ở ảnh không dẫn 703. Điều này là do việc thực hiện đó sẽ bỏ qua nửa khung được liên kết với ảnh IRAP 702. Điều này tạo ra vấn đề do các hệ thống mã hóa video sử dụng VVC có thể bị giới hạn ở vị trí các ảnh dẫn 704 ngay sau ảnh IRAP 702 theo thứ tự giải mã 708.

Sáng chế cho phép CVS 700 được sử dụng bởi hệ thống VVC. Cụ thể là, có thể được báo hiệu để chỉ báo khi một ảnh không dẫn 703 được cho phép được định vị giữa ảnh IRAP 702 và các ảnh dẫn 704. Video hệ thống có thể vẫn bị hạn chế để ngăn không cho các ảnh không dẫn 703 và/hoặc các ảnh theo sau 706 được định vị giữa các

ảnh dẫn 704. Theo đó, cò có thể chỉ báo rằng thứ tự giải mã 708 bao gồm ảnh IRAP 702, một ảnh không dẫn 703, các ảnh dẫn bất kỳ 704 (ví dụ, các ảnh dẫn 704 là tùy chọn và có thể được bỏ qua theo một số ví dụ), và sau đó bao gồm một hoặc nhiều ảnh theo sau 706. Như vậy, cò có thể chỉ báo cho bộ giải mã xem mong muốn CVS 500 hoặc CVS 700. Theo một số ví dụ, `field_seq_flag` trong SPS có thể được sử dụng nhằm mục đích như được thảo luận dưới đây.

FIG.8 là hình vẽ dạng giản lược thể hiện dòng bit làm ví dụ 800 được tạo cấu hình để chứa cả hoạt động tạo mã video đan xen và các ảnh dẫn. Ví dụ, dòng bit 800 có thể được tạo ra bởi hệ thống mã hóa 200 và/hoặc bộ mã hóa 300 để giải mã bởi hệ thống mã hóa 200 và/hoặc bộ giải mã 400 theo phương pháp 100. Ngoài ra, dòng bit 800 có thể bao gồm CVS 500 và/hoặc 700. Như vậy, dòng bit 800 có thể bao gồm ảnh thứ nhất 601 và ảnh thứ hai 602 mà có thể được kết hợp để tạo ra khung video đan xen 600. Ngoài ra, dòng bit 800 có thể bao gồm các ảnh dẫn 504 và/hoặc 704.

Dòng bit 800 bao gồm SPS 810, các tập hợp thông số ảnh (PPS) 811, các tiêu đề lát cắt 815, và dữ liệu hình ảnh 820. SPS 810 chứa chuỗi dữ liệu chung cho tất cả các ảnh trong chuỗi video đã được tạo mã có trong dòng bit 800. Dữ liệu này có thể bao gồm kích thước ảnh, chiều sâu bit, các thông số công cụ tạo mã, các hạn chế tốc độ bit, v.v. PPS 811 chứa các thông số mà áp dụng cho toàn bộ ảnh. Do đó, mỗi ảnh trong chuỗi video có thể nói đến PPS 811. Cần lưu ý rằng, trong khi mỗi ảnh là nói đến PPS 811, một PPS 811 có thể chứa dữ liệu cho nhiều ảnh theo một số ví dụ. Ví dụ, nhiều ảnh tương tự có thể được tạo mã theo các thông số tương tự. Trong trường hợp như vậy, một PPS 811 có thể chứa dữ liệu cho các ảnh tương tự này. PPS 811 có thể chỉ báo các dụng cụ tạo mã khả dụng cho các lát cắt trong các ảnh tương ứng, các thông số lượng tử hóa, các độ lệch, v.v. Tiêu đề lát 815 chứa các thông số đặc trưng cho mỗi lát cắt trong ảnh. Do đó, có thể có một tiêu đề lát 815 trên một lát cắt trong chuỗi video. Tiêu đề lát 815 có thể chứa thông tin kiểu lát cắt, các số đếm thứ tự ảnh (POC: picture order count), các danh sách ảnh tham chiếu, các trọng số dự đoán, các điểm mục nhập ảnh xếp lớp, các thông số tách khối, v.v. Cần lưu ý rằng lát cắt tiêu đề 815 cũng có thể được gọi là tiêu đề nhóm ảnh xếp lớp trong một số ngữ cảnh.

Dữ liệu hình ảnh 820 chứa dữ liệu video được mã hóa theo dự đoán liên ảnh và/hoặc dự đoán trong ảnh cũng như dữ liệu dư được chuyển đổi và được lượng tử hóa

tương ứng. Ví dụ, chuỗi video bao gồm các khung 821. Khung 821 là hình ảnh hoàn thành mà được dự định cho việc hiển thị hoàn toàn hoặc một phần đối với người dùng ở thời điểm tương ứng trong chuỗi video. Khung 821 có thể bao gồm một hoặc nhiều ảnh 823. Trong hầu hết các ngữ cảnh, khung 821 bao gồm một ảnh 823. Trong trường hợp như vậy, ảnh 823, hình ảnh/khung 821 được chứa trong một cụm truy nhập (AU). Tuy nhiên, trong ngữ cảnh video đan xen, ảnh 823 là trường của các đường nằm ngang, trường thứ nhất 610 hoặc trường thứ hai 612, chứa trong AU. Như vậy, khung 821 có thể được tạo ra từ hai ảnh 823 khi sử dụng hoạt động tạo mã video đan xen. Ảnh 823 chứa một hoặc nhiều lát cắt 825. Lát cắt 825 có thể được xác định là số nguyên các ảnh xếp lớp hoàn thành hoặc số nguyên liên tiếp các hàng cụm cây tạo mã (CTU) hoàn thành (ví dụ, bên trong ảnh xếp lớp) của ảnh 823 mà được chứa riêng trong một đơn vị lớp trừu tượng mạng (NAL: network abstraction layer). Do đó, các lát cắt 825 còn được chia thành các CTU và/hoặc các khối cây tạo mã (CTB). Các CTU/CTB còn được chia thành các khối mã hóa dựa trên các cây tạo mã. Sau đó các khối mã hóa có thể được mã hóa/được giải mã theo các cơ chế dự đoán.

Dòng bit 800 có thể bao gồm `field_seq_flag` 827. `field_seq_flag` 827 có thể được đặt tới giá trị thứ nhất khi các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự tạo mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP như được thể hiện trên CVS 500. Còn có thể được đặt tới giá trị thứ hai khi ảnh không dẫn ở trước, theo thứ tự tạo mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP và không có các ảnh dẫn được định vị giữa ảnh dẫn ban đầu và ảnh dẫn cuối cùng theo thứ tự giải mã như được thể hiện trên CVS 700. Trong trường hợp này, ảnh IRAP bao gồm trường thứ nhất của khung, và ảnh không dẫn ở trước ảnh dẫn ban đầu bao gồm trường thứ hai của khung. Theo ví dụ được thể hiện, `field_seq_flag` 827 có thể có mặt trong SPS 810. Theo ví dụ cụ thể, `field_seq_flag` 827 có thể được thiết lập to một khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh 823 mà thể hiện các trường của khung 821 hoặc được đặt thành không khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh 823 mà mỗi ảnh thể hiện khung hoàn thành 821. Theo đó, bộ giải mã có thể đọc `field_seq_flag` 827 để xác định khi việc giải mã ảnh IRAP và một hoặc nhiều ảnh không dẫn nên bao gồm việc đan xen trường thứ nhất từ ảnh IRAP và trường thứ hai từ ảnh không dẫn ở trước ảnh dẫn ban đầu để tạo ra một khung. Như vậy, `field_seq_flag` 827 cho phép hoạt động

tạo mã video đan xen được sử dụng có dựa vào các ảnh dẫn. Như vậy, việc sử dụng field_seq_flag 827 làm tăng chức năng của bộ mã hóa và/hoặc bộ giải mã. Ngoài ra, việc sử dụng field_seq_flag 827 có thể làm tăng hiệu suất tạo mã của dòng bit 800 bằng cách cho phép tốc độ khung hiệu quả tăng lên mà không cần làm tăng một cách đáng kể lượng dữ liệu mà được yêu cầu để tạo mã chuỗi video. Như vậy, sử dụng field_seq_flag 827 có thể làm giảm việc sử dụng của bộ xử lý, bộ nhớ, và/hoặc các tài nguyên truyền mạng ở bộ mã hóa và/hoặc bộ giải mã.

Thông tin nêu trên được mô tả chi tiết hơn dưới đây. Các ảnh IRAP cung cấp các chức năng có lợi khác nhau, nhưng tạo ra bất lợi cho hiệu quả nén. Sự có mặt của ảnh IRAP có thể gây ra tăng đột biến trong tốc độ bit. Điều gây bất lợi này cho hiệu quả nén có thể xuất hiện do nhiều nguyên nhân. Ví dụ, ảnh IRAP là ảnh được dự đoán trong ảnh, và do đó ảnh IRAP yêu cầu nhiều bit hơn để biểu diễn khi so sánh với các ảnh được dự đoán liên ảnh. Ngoài ra, sự có mặt của ảnh IRAP có thể phá vỡ sự dự đoán thời gian. Điều này là do bộ giải mã có thể nạp mới quy trình giải mã ngay khi nhận ảnh IRAP, mà dẫn tới việc loại bỏ các ảnh tham chiếu trước đó trong DPB. Điều này có thể khiến cho việc tạo mã của các ảnh mà theo sau ảnh IRAP theo thứ tự giải mã ít hiệu quả hơn do nó có truy nhập ít hơn tới các ảnh tham chiếu khi thực hiện việc tạo mã dự đoán liên ảnh.

Trong số các loại ảnh mà được sử dụng như các ảnh IRAP, các ảnh IDR có thể sử dụng việc báo hiệu và suy ra khác nhau khi so sánh với các loại ảnh khác. Một số khác nhau là như sau. Khi việc báo hiệu và/hoặc suy ra POC giá trị của ảnh IDR, phần bit có nghĩa tối đa (MSB) của POC có thể không được suy ra từ ảnh chính trước đó. Thay vào đó, MSB của POC có thể được đặt bằng không. Ngoài ra, lát cắt tiêu đề của ảnh IDR có thể không chứa thông tin để hỗ trợ bộ giải mã trong việc thực hiện quản lý ảnh tham chiếu. Đối các loại ảnh khác, như CRA, trailing, và thời gian lớp con truy nhập (TSA), thông tin như nhóm ảnh tham chiếu (RPS) hoặc các danh sách ảnh tham chiếu có thể có mặt trong lát cắt tiêu đề và được sử dụng cho quy trình đánh dấu các ảnh tham chiếu. Quy trình đánh dấu ảnh tham chiếu được sử dụng để xác định trạng thái của các ảnh tham chiếu trong DPB là được sử dụng cho tham chiếu hoặc không được sử dụng cho tham chiếu. Tuy nhiên, đối với các ảnh IDR, thông tin này có thể không 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.

Ngoài ra, các ảnh dẫn có thể được liên kết với ảnh IRAP. Các ảnh dẫn là các ảnh mà theo sau ảnh IRAP được kết hợp của nó theo thứ tự giải mã nhưng ở trước ảnh IRAP theo thứ tự xuất. Phụ thuộc vào cấu hình tạo mã và cấu trúc tham chiếu ảnh, các ảnh dẫn có thể còn được xác định thành hai loại. Loại thứ nhất của các ảnh, đã biết là các ảnh RASL, là các ảnh dẫn mà có thể không được giải mã một cách chính xác khi quy trình giải mã bắt đầu ở ảnh IRAP kết hợp. Điều này có thể xảy ra do này các ảnh dẫn được tạo mã tham chiếu đến các ảnh mà ở trước ảnh IRAP theo thứ tự giải mã. Loại thứ hai của các ảnh, đã biết là các ảnh RADL, là các ảnh dẫn mà có thể sẽ được giải mã một cách chính xác ngay cả khi quy trình giải mã bắt đầu ở ảnh IRAP kết hợp. Điều này là khả thi do các ảnh dẫn này được tạo mã mà không cần tham chiếu, một cách trực tiếp hoặc một cách gián tiếp, bất kỳ các ảnh ở trước ảnh IRAP theo thứ tự giải mã. Trong một số hệ thống mã hóa video, các ảnh RASL được liên kết với các ảnh IRAP được ràng buộc các ảnh RADL phía trước được liên kết với cùng một ảnh IRAP theo thứ tự xuất.

Các ảnh các ảnh IRAP và dẫn có thể được cho các loại cụm NAL khác nhau sao cho chúng có thể được nhận dạng một cách dễ dàng bởi các ứng dụng mức hệ thống. Ví dụ, bộ nói video có thể hiểu được là các loại ảnh được tạo mã mà không cần xem xét tới các phần tử cú pháp chi tiết trong dòng bit được tạo mã. Ví dụ, việc nói có thể cần xác định các ảnh IRAP từ các ảnh không IRAP và xác định các ảnh dẫn, bao gồm xác định các ảnh RASL và RADL, từ các ảnh theo sau. Các ảnh theo sau là các ảnh mà được liên kết với ảnh IRAP và theo sau ảnh IRAP theo thứ tự xuất. Ảnh hiện tại được liên kết với ảnh IRAP khi ảnh hiện tại theo sau ảnh IRAP theo thứ tự giải mã và ở trước ảnh IRAP khác bất kỳ theo thứ tự giải mã. Như vậy, việc tạo ra các ảnh IRAP và dẫn với loại cụm NAL tương ứng hỗ trợ chức năng của các ứng dụng này.

Trong một số hệ thống mã hóa video, các loại cụm NAL cho các ảnh IRAP và dẫn có thể bao gồm các thành phần sau đây. Truy nhập liên kết được phá vỡ (BLA) với ảnh dẫn (BLA_W_LP) là cụm NAL cho ảnh BLA mà có thể được theo sau bởi một hoặc nhiều ảnh dẫn theo thứ tự giải mã. BLA với RADL (BLA_W_RADL) là cụm NAL cho ảnh BLA mà có thể được theo sau bởi một hoặc nhiều ảnh RADL nhưng không có

ảnh RASL theo thứ tự giải mã. BLA với ảnh không dẫn (BLA_N_LP) là cụm NAL của ảnh BLA không được theo sau bởi ảnh dẫn theo thứ tự giải mã. IDR với RADL (IDR_W_RADL) là cụm NAL của ảnh IDR mà có thể được theo sau bởi một hoặc nhiều ảnh RADL, nhưng không có ảnh RASL, theo thứ tự giải mã. IDR với ảnh không dẫn (IDR_N_LP) là cụm NAL của ảnh IDR không được theo sau bởi ảnh dẫn theo thứ tự giải mã. CRA là cụm NAL của ảnh CRA mà có thể được theo sau bởi các ảnh dẫn bao gồm các ảnh RASL và/hoặc RADL. RADL là cụm NAL của ảnh RADL. RASL là cụm NAL của ảnh RASL.

Các hệ thống mã hóa video khác có thể sử dụng các loại cụm NAL sau đây cho các ảnh IRAP và dẫn. IDR_W_RADL là cụm NAL của ảnh IDR mà có thể được theo sau bởi một hoặc nhiều ảnh RADL, nhưng không có ảnh RASL, theo thứ tự giải mã. IDR_N_LP là cụm NAL của ảnh IDR không được theo sau bởi ảnh dẫn theo thứ tự giải mã. CRA là cụm NAL của ảnh CRA mà có thể được theo sau bởi các ảnh dẫn như các ảnh RASL và/hoặc RADL. RADL là cụm NAL của ảnh RADL. RASL là cụm NAL của ảnh RASL.

Đối với độ thích ứng dòng bit, một số ràng buộc có thể được áp dụng vào các ảnh dẫn, ví dụ trong HEVC và/hoặc các hệ thống VVC. Các ràng buộc này là như sau. Mỗi ảnh, khác với ảnh thứ nhất trong dòng bit theo thứ tự giải mã, có thể được xem là được liên kết với ảnh IRAP trước đó theo thứ tự giải mã. Khi ảnh là ảnh dẫn của ảnh IRAP, ảnh sẽ là ảnh RADL hoặc RASL. Khi ảnh là ảnh theo sau của ảnh IRAP, ảnh sẽ không là ảnh RADL hoặc RASL. Khi ảnh là ảnh dẫn của ảnh IRAP, ảnh sẽ ở trước, theo thứ tự giải mã, tất cả các ảnh theo sau mà được liên kết với cùng một ảnh IRAP. Không có các ảnh RASL được liên kết với ảnh IDR. Không có các ảnh RADL shall be được liên kết với ảnh IDR có nal_unit_type bằng với IDR_N_LP. Cần lưu ý rằng truy nhập ngẫu nhiên có thể được thực hiện ở vị trí của cụm truy nhập IRAP bằng cách hủy bỏ tất cả các cụm truy nhập trước cụm truy nhập IRAP. Sự truy nhập ngẫu nhiên như vậy có thể dẫn tới việc giải mã một cách chính xác ảnh IRAP và tất cả các ảnh không RASL tiếp theo theo thứ tự giải mã. Sự truy nhập ngẫu nhiên như vậy có thể được thực hiện, với điều kiện là mỗi tập thông số là khả dụng, trong dòng bit hoặc bởi phương tiện bên ngoài như đầu vào của người dùng, khi tập thông số này được kích hoạt. Ngoài ra, ảnh bất kỳ mà ở trước ảnh IRAP theo thứ tự giải mã sẽ ở trước ảnh IRAP theo thứ tự xuất

và sẽ ở trước ảnh RADL bất kỳ được liên kết với ảnh IRAP theo thứ tự xuất. Ảnh RASL bất kỳ được liên kết với ảnh CRA sẽ ở trước ảnh RADL bất kỳ được liên kết với ảnh CRA theo thứ tự xuất. Ảnh RASL bất kỳ được liên kết với ảnh CRA sẽ theo sau, theo thứ tự xuất, ảnh IRAP bất kỳ mà ở trước ảnh CRA theo thứ tự giải mã.

Theo đó, các ràng buộc thích ứng dòng bit liên quan đến các ảnh dẫn như được mô tả trên đây có thể xung đột với các cơ chế tạo mã video đan xen. Sự xung đột là như sau. Khi hoạt động tạo mã đan xen được sử dụng, hai trường của ảnh IRAP đều không được đánh dấu là các ảnh IRAP. Thay vào đó, chỉ trường thứ nhất được đánh dấu là ảnh IRAP và trường thứ hai được đánh dấu là ảnh theo sau. Ảnh theo sau được đan xen, mà chứa trường thứ hai của ảnh, nên theo sau ngay ảnh IRAP được đan xen theo thứ tự giải mã. Điều này là do ảnh IRAP được đan xen và ảnh được đan xen theo sau tạo ra khung hoàn thành. Nếu (các) ảnh dẫn theo sau ảnh IRAP theo thứ tự giải mã, sau đó ràng buộc biểu thị rằng khi ảnh là ảnh dẫn của ảnh IRAP, ảnh sẽ ở trước, theo thứ tự giải mã, tất cả các ảnh theo sau mà được liên kết với cùng một ảnh IRAP. Ràng buộc nêu trên không thể được loại bỏ một cách đơn giản do ràng buộc có thể giúp thực thể bên ngoài, như bộ nối video, xác định một cách hiệu quả xem có ảnh dẫn được liên kết với IRAP và liệu tất cả các ảnh dẫn có được xem xét hay không. Thực thể bên ngoài này có thể hoạt động như sau. Bắt đầu từ ảnh IRAP, nếu ảnh mà theo ngay sau ảnh IRAP là ảnh theo sau, thực thể bên ngoài có thể xác định rằng không có các ảnh dẫn được liên kết với ảnh IRAP. Do đó, để tìm kiếm tất cả các ảnh dẫn được liên kết với ảnh IRAP, thực thể bên ngoài có thể tìm ảnh theo sau thứ nhất mà theo sau ảnh IRAP theo thứ tự giải mã dựa trên ràng buộc này. Không cần ràng buộc nêu trên, thực thể bên ngoài có thể được yêu cầu tìm kiếm cho đến khi ảnh IRAP tiếp theo để tìm ra tất cả các ảnh dẫn được liên kết với ảnh IRAP.

Nói chung, sáng chế mô tả các phương pháp để xử lý các ảnh dẫn được liên kết với các ảnh IRAP. Cụ thể hơn là, sáng chế mô tả các phương pháp để tìm kiếm và xác định một cách hiệu quả các ảnh dẫn được liên kết với ảnh IRAP trong khi hỗ trợ hoạt động tạo mã hiệu quả của nội dung video xen kẽ. Phần mô tả của các công nghệ được mô tả dựa trên VVC tiêu chuẩn bởi JVET của ITU-T và ISO/IEC. Tuy nhiên, các công nghệ cũng có thể áp dụng vào các đặc tả mã hóa video khác.

Để giải quyết các vấn đề được liệt kê bên trên, sáng chế bao gồm các khía cạnh sau đây, mà có thể được áp dụng riêng biệt hoặc theo cách kết hợp. Ví dụ, các ảnh dẫn được liên kết với ảnh IRAP có thể được định vị một cách liên tục theo thứ tự giải mã mà không có (các) ảnh không dẫn ở giữa. Ngoài ra, ràng buộc sau đây được áp dụng cho sự tương thích dòng bit của các ảnh IRAP và dẫn. Lấy picA và picB là các ảnh dẫn đầu tiên và cuối cùng được liên kết với ảnh IRAP, một cách tương ứng. Trong trường hợp như vậy, sẽ không có ảnh mà không là ảnh dẫn theo sau picA và ở trước picB theo thứ tự giải mã, một cách tương ứng.

Các ràng buộc sau đây cũng có thể được áp dụng. Nếu `field_seq_flag` được chọn bằng không và ảnh hiện tại là ảnh dẫn được liên kết với ảnh IRAP, ảnh hiện tại sẽ ở trước, theo thứ tự giải mã, tất cả các ảnh không dẫn mà được liên kết với cùng một ảnh IRAP. Ngược lại, nếu `field_seq_flag` được chọn bằng một, lấy picA và picB là các ảnh dẫn đầu tiên và cuối cùng, theo thứ tự giải mã, được liên kết với ảnh IRAP, một cách tương ứng. Trong trường hợp như vậy, sẽ có nhiều nhất một ảnh không dẫn ở trước picA theo thứ tự giải mã, và sẽ không có ảnh không dẫn mà kế tiếp picA theo thứ tự giải mã và ở trước picB theo thứ tự giải mã.

Các ràng buộc sau đây cũng có thể được áp dụng. Nếu `general_frame_only_constraint_flag` bằng một và ảnh hiện tại là ảnh dẫn được liên kết với ảnh IRAP, ảnh hiện tại sẽ ở trước, theo thứ tự giải mã, tất cả các ảnh không dẫn mà được liên kết với cùng một ảnh IRAP. Ngược lại, nếu `general_frame_only_constraint_flag` bằng một, lấy picA và picB là các ảnh dẫn đầu tiên và cuối cùng, theo thứ tự giải mã, được liên kết với ảnh IRAP, một cách tương ứng. Trong trường hợp như vậy, sẽ có nhiều nhất một ảnh không dẫn ở trước picA theo thứ tự giải mã, và sẽ không có ảnh không dẫn mà kế tiếp picA theo thứ tự giải mã và ở trước picB theo thứ tự giải mã.

Theo một ví dụ, loại cụm NAL của ảnh IRAP cung cấp thông tin thích hợp để xác định xem (các) ảnh hàng đầu được liên kết với ảnh IRAP có mặt hay không. Để thực hiện điều này, phương pháp sau đây có thể được sử dụng. NAL cụm loại `CRA_NUT` có thể được thay thế bằng `CRA_W_LP` để chỉ báo rằng (các) ảnh dẫn được liên kết với ảnh CRA và/hoặc `CRA_N_LP` để chỉ báo không có các ảnh dẫn được liên kết với ảnh CRA. Theo một ví dụ khác, các loại cụm NAL `IDR_W_RADL`, `IDR_N_LP`, và `CRA_NUT` có thể được thay thế bằng `IRAP_W_LP` để chỉ báo rằng các ảnh dẫn

được liên kết với ảnh IRAP và IRAP_N_LP để chỉ báo rằng các ảnh dẫn không được liên kết với ảnh IRAP.

Theo một ví dụ, dưới đây có thể áp dụng cho CRA_W_LP, CRA_N_LP, IDR_W_RADL, và IDR_N_LP. Ảnh IDR có NalUnitType bằng IDR_N_LP không được liên kết với các ảnh dẫn bất kỳ có trong dòng bit. Ảnh IDR có NalUnitType bằng IDR_W_RADL không được liên kết với các ảnh RASL có trong dòng bit, nhưng có thể được liên kết với các ảnh RADL trong dòng bit. Ảnh CRA có NalUnitType bằng CRA_N_LP không được liên kết với các ảnh dẫn có trong dòng bit. Ảnh CRA có NalUnitType bằng CRA_W_LP có thể được liên kết với các ảnh dẫn trong dòng bit.

Theo một ví dụ, ánh xạ của các loại cụm NAL nêu trên tới các loại điểm truy nhập luồng (SAP) là như sau. IDR_N_LP và CRA_N_LP được liên kết với SAP loại 1, IDR_W_RADL được liên kết với SAP loại 2, và CRA_W_LP được liên kết với SAP loại 3.

Theo một ví dụ, phần dưới đây có thể áp dụng cho IRAP_W_LP và IRAP_N_LP. Ảnh IRAP có NalUnitType bằng IRAP_N_LP không kết hợp với các ảnh dẫn có trong dòng bit. Ảnh IRAP có NalUnitType bằng IRAP_W_LP có thể được liên kết với các ảnh dẫn trong dòng bit.

Theo một ví dụ, ánh xạ của các loại cụm NAL nêu trên tới các loại SAP là như sau. IRAP_N_LP được liên kết với SAP loại 1 và IRAP_W_LP được liên kết với SAP loại 3.

Theo một ví dụ, để xác định xem các ảnh dẫn được liên kết với IRAP có mặt hay không, thiết bị có thể kiểm tra loại cụm NAL của ảnh IRAP. Khi ảnh IRAP có thể được liên kết với một hoặc nhiều ảnh dẫn được kết hợp, bước sau đây có thể được sử dụng để tìm tất cả các ảnh dẫn được liên kết với ảnh IRAP. Thiết bị có thể bắt đầu từ ảnh IRAP. Nếu ảnh mà theo ngay sau ảnh IRAP theo thứ tự giải mã là ảnh không dẫn, ảnh này có thể được bỏ qua. Lưu ý rằng sự có mặt của ảnh không dẫn này ngay sau khi ảnh IRAP có thể chỉ báo rằng dòng bit là dòng bit tạo mã video đan xen. Hình ảnh tiếp theo sẽ là ảnh dẫn. Quy trình có thể tiếp tục kiểm tra hình ảnh tiếp theo cho đến khi gặp ảnh không dẫn thứ nhất.

FIG.9 là hình vẽ dạng sơ đồ của thiết bị mã hóa video làm ví dụ 900. Thiết bị mã hóa video 900 có thể thích hợp để thực hiện các ví dụ/các phương án thực hiện được

bộc lộ như đã được mô tả ở đây. Thiết bị mã hóa video 900 bao gồm cổng phía dưới 920, cổng phía trên 950, và/hoặc bộ thu phát đơn vị (Tx/Rx) 910, bao gồm các bộ truyền và/hoặc các bộ thu để truyền thông dữ liệu phía trên và/hoặc phía dưới trên mạng. Thiết bị mã hóa video 900 còn bao gồm bộ xử lý 930 bao gồm cụm logic và/hoặc bộ phận xử lý trung tâm (CPU) để xử lý dữ liệu và bộ nhớ 932 để lưu trữ dữ liệu. Thiết bị mã hóa video 900 có thể còn bao gồm các thành phần điện, quang điện (OE: optical-to-electrical), các thành phần điện quang (EO: electrical-to-optical), và/hoặc các thành phần truyền thông không dây ghép nối với cổng phía trên 950 và/hoặc cổng phía dưới 920 để truyền thông dữ liệu thông qua các mạng truyền thông bằng điện, quang hoặc không dây. Thiết bị mã hóa video 900 có thể còn bao gồm các thiết bị vào và/hoặc ra (I/O) 960 để truyền thông dữ liệu đến và từ người dùng. Các thiết bị I/O 960 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 960 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 xử lý 930 được thực hiện bởi phần cứng và phần mềm. Bộ xử lý 930 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 tắt 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 930 truyền thông với các cổng cuối dòng 920, Tx/Rx 910, các cổng đầu dòng 950, và bộ nhớ 932. Bộ xử lý 930 bao gồm môđun tạo mã 914. Môđun mã hóa 914 thực hiện các phương án được bộc lộ đã được mô tả ở đây, như các phương pháp 100, 1000, và 1100, mà có thể sử dụng CVS 500, khung video đan xen 600, CVS 700, và/hoặc dòng bit 800. Môđun mã hóa 914 cũng có thể thực hiện phương pháp/cơ chế khác bất kỳ đã được mô tả ở đây. Ngoài ra, môđun mã hóa 914 có thể thực hiện hệ thống mã hóa 200, bộ mã hóa 300, và/hoặc bộ giải mã 400. Ví dụ, môđun mã hóa 914 có thể đặt cờ trong SPS để chỉ báo khi ảnh không dẫn được định vị giữa ảnh IRAP và tập hợp của các ảnh dẫn. Do đó, môđun mã hóa 914 khiến cho thiết bị mã hóa video 900 cung cấp bổ sung chức năng và/hoặc hiệu suất tạo mã khi tạo mã dữ liệu video. Như vậy, môđun mã hóa 914 cải thiện chức năng của thiết bị mã hóa video 900 cũng như khắc phục các vấn

đề mà đặc trưng cho lĩnh vực tạo mã video. Ngoài ra, môđun mã hóa 914 có tác dụng chuyển đổi thiết bị mã hóa video 900 thành trạng thái khác nhau. Theo cách thay thế, môđun mã hóa 914 có thể được thực hiện như các lệnh chứa trong bộ nhớ 932 và được thực thi bởi bộ xử lý 930 (ví dụ, làm sản phẩm chương trình máy tính được lưu trữ trên phương tiện lâu dài).

Bộ nhớ 932 bao gồm một hoặc nhiều loại bộ nhớ như các đĩa, các ổ băng, các ổ đĩa thể rắn, bộ nhớ chỉ đọc (read only memory, viết tắt là ROM), bộ nhớ truy cập ngẫu nhiên (random access memory, viết tắt là RAM), bộ nhớ chớp, bộ nhớ có thể nhắm đến nội dung tam phân (ternary content-addressable memory, viết tắt là TCAM), bộ nhớ truy cập ngẫu nhiên tĩnh (static random-access memory viết tắt là SRAM), v.v.. Bộ nhớ 932 có thể được sử dụng làm thiết bị lưu trữ dữ liệu tràn, để lưu trữ các chương trình khi các chương trình này được lựa chọn thực hiện, và lưu trữ các chỉ dẫn và dữ liệu mà được đọc trong khi thực hiện chương trình.

FIG.10 là lưu đồ của phương pháp làm ví dụ 1000 của mã hóa chuỗi video, như CVS 500 và/hoặc 700, với hoạt động tạo mã video đan xen, như khung video đan xen 600, và các ảnh dẫn vào trong dòng bit, như dòng bit 800. Phương pháp 1000 có thể được sử dụng bởi bộ mã hóa, như hệ thống mã hóa 200, bộ mã hóa 300, và/hoặc thiết bị mã hóa video 900 khi thực hiện phương pháp 100.

Phương pháp 1000 có thể bắt đầu khi bộ mã hóa nhận chuỗi video bao gồm các ảnh và xác định để mã hóa chuỗi video đó vào trong dòng bit, ví dụ dựa trên đầu vào của người dùng. Ở bước 1001, bộ mã hóa xác định thứ tự tạo mã cho chuỗi video. Chuỗi video bao gồm các ảnh có ảnh IRAP và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP. Theo cách tùy chọn chuỗi video cũng có thể bao gồm một hoặc nhiều (ví dụ, nhóm) ảnh trong số các ảnh dẫn.

Ở bước 1003, bộ mã hóa có thể mã hóa cờ vào trong dòng bit. Cờ có thể được đặt tới giá trị thứ nhất khi các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự tạo mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP như trong CVS 500. Điều này chỉ báo rằng chuỗi video không chứa video đan xen. Cờ cũng có thể được đặt tới giá trị thứ hai khi ảnh không dẫn ở trước, theo thứ tự tạo mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP như trong CVS 700. Khi cờ được đặt thành giá trị thứ hai, dòng bit cũng có thể được ràng buộc sao cho không có các ảnh dẫn được định vị giữa ảnh dẫn

ban đầu và ảnh dẫn cuối cùng theo thứ tự tạo mã. Điều này có thể chỉ báo rằng chuỗi video chứa video đan xen. Theo ví dụ cụ thể, bộ mã hóa có thể mã hóa SPS vào trong dòng bit, và cờ có thể được mã hóa vào trong SPS. Theo một số ví dụ, cờ là `field_seq_flag`. Ví dụ, `field_seq_flag` có thể được đặt thành một khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường. Ngoài ra, `field_seq_flag` có thể được đặt thành không khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung. Theo đó, cờ có thể được đặt để chỉ báo rằng hoạt động tạo mã video đan xen được sử dụng trong dòng bit. Như vậy, cờ có thể được đặt khi ảnh IRAP bao gồm trường thứ nhất của khung, và ảnh không dẫn ở trước ảnh dẫn ban đầu bao gồm trường thứ hai của khung. Ví dụ, trường thứ nhất từ ảnh IRAP và trường thứ hai từ ảnh không dẫn ở trước ảnh dẫn ban đầu có thể bao gồm việc xen kẽ các đường của dữ liệu video mà thể hiện một khung video đan xen như được thể hiện đối với FIG.6A đến Fig.6C .

Ở bước 1005, bộ mã hóa có thể mã hóa ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP vào trong dòng bit theo thứ tự tạo mã. Sau đó, bộ mã hóa có thể lưu trữ dòng bit để truyền thông về phía bộ giải mã ở bước 1007.

FIG.11 là lưu đồ của phương pháp làm ví dụ 1100 để giải mã chuỗi video, như CVS 500 và/hoặc 700, với hoạt động tạo mã video đan xen, như khung video đan xen 600, và các ảnh dẫn từ dòng bit, như dòng bit 800. Phương pháp 1100 có thể được sử dụng bởi bộ giải mã, như hệ thống mã hóa 200, bộ giải mã 400, và/hoặc thiết bị mã hóa video 900 khi thực hiện phương pháp 100.

Phương pháp 1100 có thể bắt đầu khi bộ giải mã bắt đầu nhận dòng bit của dữ liệu được tạo mã đại diện cho chuỗi video, ví dụ nhờ phương pháp 1000. Ở bước 1101, bộ giải mã nhận dòng bit bao gồm cờ và các ảnh được tạo mã bao gồm ảnh IRAP và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP. Theo cách tùy chọn chuỗi video cũng có thể bao gồm một hoặc nhiều (ví dụ, nhóm) ảnh trong số các ảnh dẫn.

Ở bước 1103, bộ giải mã có thể xác định rằng các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự giải mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP khi cờ được đặt thành giá trị thứ nhất; Điều này chỉ báo rằng chuỗi video không chứa video đan xen. Ở bước 1105, bộ giải mã có thể xác định rằng ảnh không

dẫn ở trước, theo thứ tự giải mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP khi cờ được đặt thành giá trị thứ hai như được thể hiện trên CVS 700. Khi cờ được đặt thành giá trị thứ hai bộ giải mã có thể còn xác định rằng không có các ảnh dẫn được định vị giữa ảnh dẫn ban đầu và ảnh dẫn cuối cùng theo thứ tự tạo mã. Điều này có thể chỉ báo rằng chuỗi video chứa video đan xen. Theo ví dụ cụ thể, dòng bit có thể bao gồm SPS, và cờ có thể thu được từ SPS. Theo một số ví dụ, cờ là `field_seq_flag`. Ví dụ, `field_seq_flag` có thể được đặt thành một khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường. Ngoài ra, `field_seq_flag` có thể được đặt thành không khi chỉ báo rằng chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung. Theo đó, cờ có thể được đặt để chỉ báo rằng hoạt động tạo mã video đan xen được sử dụng trong dòng bit. Như vậy, cờ có thể được đặt khi ảnh IRAP bao gồm trường thứ nhất của khung, và ảnh không dẫn ở trước ảnh dẫn ban đầu bao gồm trường thứ hai của khung.

Ở bước 1107, bộ giải mã giải mã ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP theo thứ tự giải mã dựa trên cờ. Ví dụ, việc giải mã ảnh IRAP, các ảnh dẫn (nếu có), và một hoặc nhiều ảnh không dẫn có thể bao gồm đan xen trường thứ nhất từ ảnh IRAP và trường thứ hai từ ảnh không dẫn ở trước ảnh dẫn ban đầu để tạo ra một khung như được thể hiện đối với FIG.6A đến FIG.6C. Ở bước 1109, bộ giải mã có thể chuyển tiếp một hoặc nhiều ảnh đã được giải mã là kết quả của bước 1107 để hiển thị làm một phần của chuỗi video đã được giải mã.

FIG.12 là hình vẽ dạng giản lược của hệ thống làm ví dụ 1200 để tạo mã chuỗi video, như CVS 500 và/hoặc CVS 700, với hoạt động tạo mã video đan xen, như khung video đan xen 600, và các ảnh dẫn vào trong dòng bit, như dòng bit 800. Hệ thống 1200 có thể được thực hiện bởi bộ mã hóa và bộ giải mã như hệ thống mã hóa 200, bộ mã hóa 300, bộ giải mã 400, và/hoặc thiết bị mã hóa video 900. Ngoài ra, hệ thống 1200 có thể được sử dụng khi thực hiện phương pháp 100, 1000, và/hoặc 1100.

Hệ thống 1200 bao gồm bộ mã hóa video 1202. Bộ mã hóa video 1202 bao gồm môđun xác định 1201 để xác định thứ tự tạo mã cho chuỗi video bao gồm các ảnh có ảnh IRAP và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP. Bộ mã hóa video 1202 còn bao gồm môđun mã hóa 1203 để mã hóa cờ vào trong dòng bit, trong

đó cờ được đặt thành giá trị thứ nhất khi các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự tạo mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP, và trong đó cờ được đặt thành giá trị thứ hai khi ảnh không dẫn ở trước, theo thứ tự tạo mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP. Môđun mã hóa 1203 còn để mã hóa ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP vào trong dòng bit theo thứ tự tạo mã. Bộ mã hóa video 1202 còn bao gồm môđun lưu trữ 1205 để lưu trữ dòng bit để truyền thông về phía bộ giải mã. Bộ mã hóa video 1202 còn bao gồm môđun truyền 1207 để truyền dòng bit về phía bộ giải mã video 1210. Bộ mã hóa video 1202 có thể còn được tạo cấu hình để thực hiện bước bất kỳ trong số các bước của phương pháp 1000.

Hệ thống 1200 cũng bao gồm bộ giải mã video 1210. Bộ giải mã video 1210 bao gồm môđun nhận 1211 để nhận dòng bit bao gồm cờ và các ảnh được tạo mã bao gồm ảnh IRAP và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP. Bộ giải mã video 1210 còn bao gồm môđun xác định 1213 để xác định các ảnh dẫn bất kỳ được liên kết với ảnh IRAP ở trước, theo thứ tự giải mã, tất cả các ảnh không dẫn được liên kết với ảnh IRAP khi cờ được đặt thành giá trị thứ nhất. Môđun xác định 1213 còn để xác định ảnh không dẫn ở trước, theo thứ tự giải mã, ảnh dẫn ban đầu được liên kết với ảnh IRAP khi cờ được đặt thành giá trị thứ hai. Bộ giải mã video 1210 còn bao gồm môđun giải mã 1215 để giải mã ảnh IRAP, các ảnh dẫn bất kỳ được liên kết với ảnh IRAP, và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP theo thứ tự giải mã dựa trên cờ. Bộ giải mã video 1210 còn bao gồm môđun chuyển tiếp 1217 để chuyển tiếp một hoặc nhiều ảnh đã được giải mã để hiển thị làm một phần của chuỗi video đã được giải mã. Bộ giải mã video 1210 có thể còn được tạo cấu hình để thực hiện bước bất kỳ trong số các bước của phương pháp 1100.

Thành phần thứ nhất được ghép nối trực tiếp với thành phần thứ hai khi không có các thành phần xen kẽ, ngoại trừ đường dây, đường cáp, hoặc phương tiện khác giữa thành phần thứ nhất và thành phần thứ hai. Thành phần thứ nhất được ghép nối gián tiếp với thành phần thứ hai khi có các thành phần xen kẽ khác ngoài đường dây, đường cáp, hoặc phương tiện khác giữa thành phần thứ nhất và thành phần thứ hai. Thuật ngữ “được ghép nối” và các biến thể của thuật ngữ này bao gồm cả được ghép nối trực tiếp và được

ghép nối gián tiếp. Việc sử dụng thuật ngữ “khoảng” nghĩa là phạm vi bao gồm $\pm 10\%$ của số tiếp theo trừ khi được thể hiện theo cách khác.

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 có thể được hiểu rằng các hệ thống và các phương pháp đã được bộc lộ có thể được thể hiện theo nhiều dạng cụ thể khác mà không trệt 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 thành phần, các kỹ thuật, hoặc các phương pháp khác mà không trệt khỏi phạm vi của sáng chế. 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 trệt khỏi nguyên lý và phạm vi được bộc lộ ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã được thực hiện ở bộ giải mã, trong đó phương pháp này bao gồm các bước:

nhận, bởi bộ nhận của bộ giải mã, dòng bit bao gồm cờ trường tuần tự, `field_seq_flag`, và các ảnh được tạo mã bao gồm ảnh điểm truy nhập ngẫu nhiên trong ảnh (Intra Random Access Point - IRAP) và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP này;

xác định, bởi bộ xử lý của bộ giải mã, rằng các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này là sẽ đi trước, theo thứ tự giải mã, tất cả các ảnh không dẫn mà được liên kết với ảnh IRAP này, khi `field_seq_flag` được đặt bằng 0;

xác định, bởi bộ xử lý, rằng nhiều nhất là một ảnh không dẫn đi trước, theo thứ tự giải mã, ảnh dẫn ban đầu mà được liên kết với ảnh IRAP này, và không có ảnh không dẫn nào được định vị giữa ảnh dẫn ban đầu này và ảnh dẫn cuối cùng theo thứ tự giải mã khi `field_seq_flag` được đặt bằng 1;

trong đó `field_seq_flag` được đặt bằng 1 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường, và trong đó `field_seq_flag` được đặt bằng 0 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung; và

áp dụng, bởi bộ xử lý, quy trình giải mã cho ảnh IRAP, các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này và một hoặc nhiều ảnh không dẫn mà được liên kết với ảnh IRAP này theo thứ tự giải mã, dựa trên việc `field_seq_flag` được đặt bằng 1 hay 0.

2. Phương pháp theo điểm 1, trong đó dòng bit bao gồm tập thông số chuỗi (Sequence Parameter Set - SPS), và trong đó cờ nêu trên là được thu nhận từ SPS này.

3. Phương pháp theo điểm 1 hoặc 2, trong đó ảnh IRAP bao gồm trường thứ nhất của khung, và trong đó ảnh không dẫn mà đi trước ảnh dẫn ban đầu thì bao gồm trường thứ hai của khung này.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó bước áp dụng quy trình giải mã cho ảnh IRAP và một hoặc nhiều ảnh không dẫn là bao gồm bước đan xen trường thứ nhất từ ảnh IRAP và trường thứ hai từ ảnh không dẫn mà đi trước ảnh dẫn ban đầu để tạo ra một khung.

5. Phương pháp mã hóa được thực hiện ở bộ mã hóa, trong đó phương pháp này bao gồm các bước:

xác định, bởi bộ xử lý của bộ mã hóa, thứ tự tạo mã cho chuỗi video bao gồm các ảnh mà bao gồm ảnh điểm truy nhập ngẫu nhiên trong ảnh (Intra Random Access Point - IRAP) và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP này;

mã hóa, bởi bộ xử lý, cờ trường tuần tự, `field_seq_flag`, vào dòng bit, trong đó `field_seq_flag` này được đặt bằng 0 khi các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này đi trước, theo thứ tự tạo mã, tất cả các ảnh không dẫn mà được liên kết với ảnh IRAP này, và trong đó `field_seq_flag` này được đặt bằng 1 khi nhiều nhất là một ảnh không dẫn đi trước, theo thứ tự tạo mã, ảnh dẫn ban đầu mà được liên kết với ảnh IRAP này, và không có ảnh không dẫn nào được định vị giữa ảnh dẫn ban đầu này và ảnh dẫn cuối cùng theo thứ tự giải mã;

trong đó `field_seq_flag` được đặt bằng 1 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường, và trong đó `field_seq_flag` được đặt bằng 0 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung; và

mã hóa, bởi bộ xử lý, ảnh IRAP này, các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này và một hoặc nhiều ảnh không dẫn mà được liên kết với ảnh IRAP này vào dòng bit theo thứ tự tạo mã.

6. Phương pháp theo điểm 5, trong đó dòng bit bao gồm tập thông số chuỗi (Sequence Parameter Set - SPS), và trong đó cờ nêu trên được mã hóa vào SPS này.

7. Phương pháp theo điểm 5 hoặc 6, trong đó ảnh IRAP bao gồm trường thứ nhất của khung, và trong đó ảnh không dẫn mà đi trước ảnh dẫn ban đầu thì bao gồm trường thứ hai của khung này.

8. Phương pháp theo điểm bất kỳ trong số các điểm từ 5 đến 7, trong đó trường thứ nhất từ ảnh IRAP và trường thứ hai từ ảnh không dẫn mà đi trước ảnh dẫn ban đầu là bao gồm các đường xen kẽ của một khung video được đan xen.

9. Phương tiện đọc được bằng máy tính bất biến bao gồm sản phẩm chương trình máy tính để sử dụng bởi thiết bị mã hóa video, sản phẩm chương trình máy tính này bao gồm các lệnh có thể thực thi bằng máy tính được lưu trữ trên phương tiện đọc được bằng máy tính bất biến này sao cho khi được thực thi bởi bộ xử lý thì khiến cho thiết bị tạo mã video thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 4.

10. Phương tiện đọc được bằng máy tính bất biến bao gồm sản phẩm chương trình máy tính để sử dụng bởi thiết bị mã hóa video, sản phẩm chương trình máy tính này bao gồm các lệnh có thể thực thi bằng máy tính được lưu trữ trên phương tiện đọc được bằng máy tính bất biến này sao cho khi được thực thi bởi bộ xử lý thì khiến cho thiết bị tạo mã video thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 5 đến 8.

11. Bộ giải mã bao gồm:

phương tiện nhận để nhận dòng bit bao gồm cờ trường tuần tự, `field_seq_flag`, và các ảnh được tạo mã bao gồm ảnh điểm truy nhập ngẫu nhiên trong ảnh (Intra Random Access Point - IRAP) và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP này;

phương tiện xác định để:

xác định rằng các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này là sẽ đi trước, theo thứ tự giải mã, tất cả các ảnh không dẫn mà được liên kết với ảnh IRAP này, khi `field_seq_flag` được đặt bằng 0; và

xác định rằng nhiều nhất là một ảnh không dẫn đi trước, theo thứ tự giải mã, ảnh dẫn ban đầu mà được liên kết với ảnh IRAP này, và không có ảnh không dẫn nào được định vị giữa ảnh dẫn ban đầu này và ảnh dẫn cuối cùng theo thứ tự giải mã khi `field_seq_flag` được đặt bằng 1;

trong đó `field_seq_flag` được đặt bằng 1 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường, và trong đó `field_seq_flag` được đặt bằng 0 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung;

phương tiện giải mã để áp dụng quy trình giải mã cho ảnh IRAP, các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này và một hoặc nhiều ảnh không dẫn mà được liên kết với ảnh IRAP này theo thứ tự giải mã, dựa trên việc `field_seq_flag` được đặt bằng 1 hay 0.

12. Bộ giải mã theo điểm 11, trong đó bộ giải mã này còn được tạo cấu hình để thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 2 đến 4.

13. Bộ mã hóa bao gồm:

phương tiện xác định để xác định thứ tự tạo mã cho chuỗi video bao gồm các ảnh mà bao gồm ảnh điểm truy nhập ngẫu nhiên trong ảnh (Intra Random Access Point - IRAP) và một hoặc nhiều ảnh không dẫn được liên kết với ảnh IRAP này;

phương tiện mã hóa để:

mã hóa cờ trường tuần tự, `field_seq_flag`, vào dòng bit, trong đó `field_seq_flag` này được đặt bằng 0 khi các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này đi trước, theo thứ tự tạo mã, tất cả các ảnh không dẫn mà được liên kết với ảnh IRAP này, và trong đó `field_seq_flag` này được đặt bằng 1 khi nhiều nhất là một ảnh không dẫn đi trước, theo thứ tự tạo mã, ảnh dẫn ban đầu mà được liên kết với ảnh IRAP này, và không có ảnh không dẫn nào được định vị giữa ảnh dẫn ban đầu này và ảnh dẫn cuối cùng theo thứ tự giải mã;

trong đó `field_seq_flag` được đặt bằng 1 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các trường, và trong đó `field_seq_flag` được đặt bằng 0 khi chuỗi video đã được tạo mã bao gồm các ảnh mà thể hiện các khung; và

mã hóa ảnh IRAP này, các ảnh dẫn bất kỳ mà được liên kết với ảnh IRAP này và một hoặc nhiều ảnh không dẫn mà được liên kết với ảnh IRAP này vào dòng bit theo thứ tự tạo mã.

14. Bộ mã hóa theo điểm 13, trong đó bộ mã hóa này còn được tạo cấu hình để thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 6 đến 8.

1/12

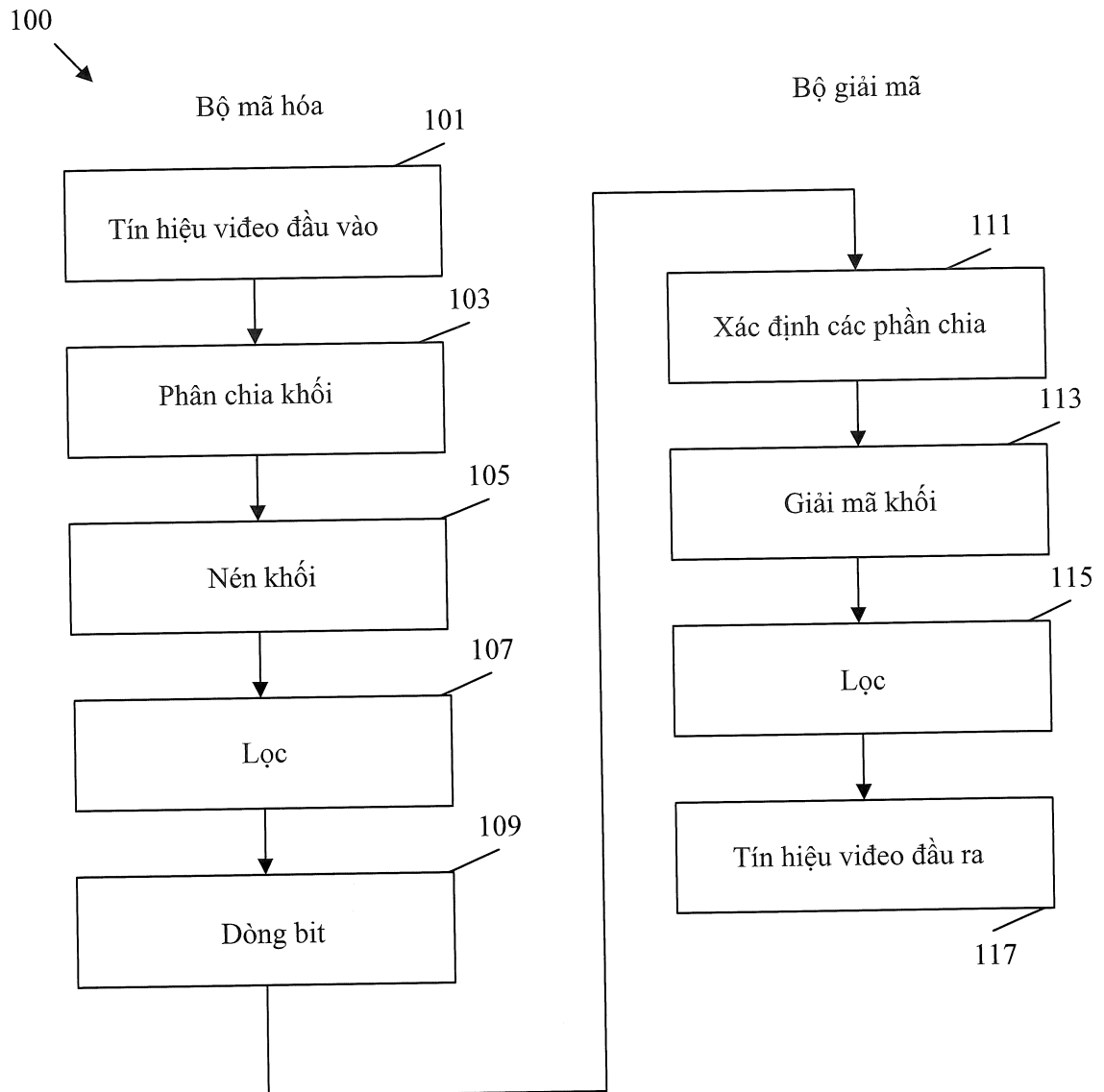


FIG. 1

2/12

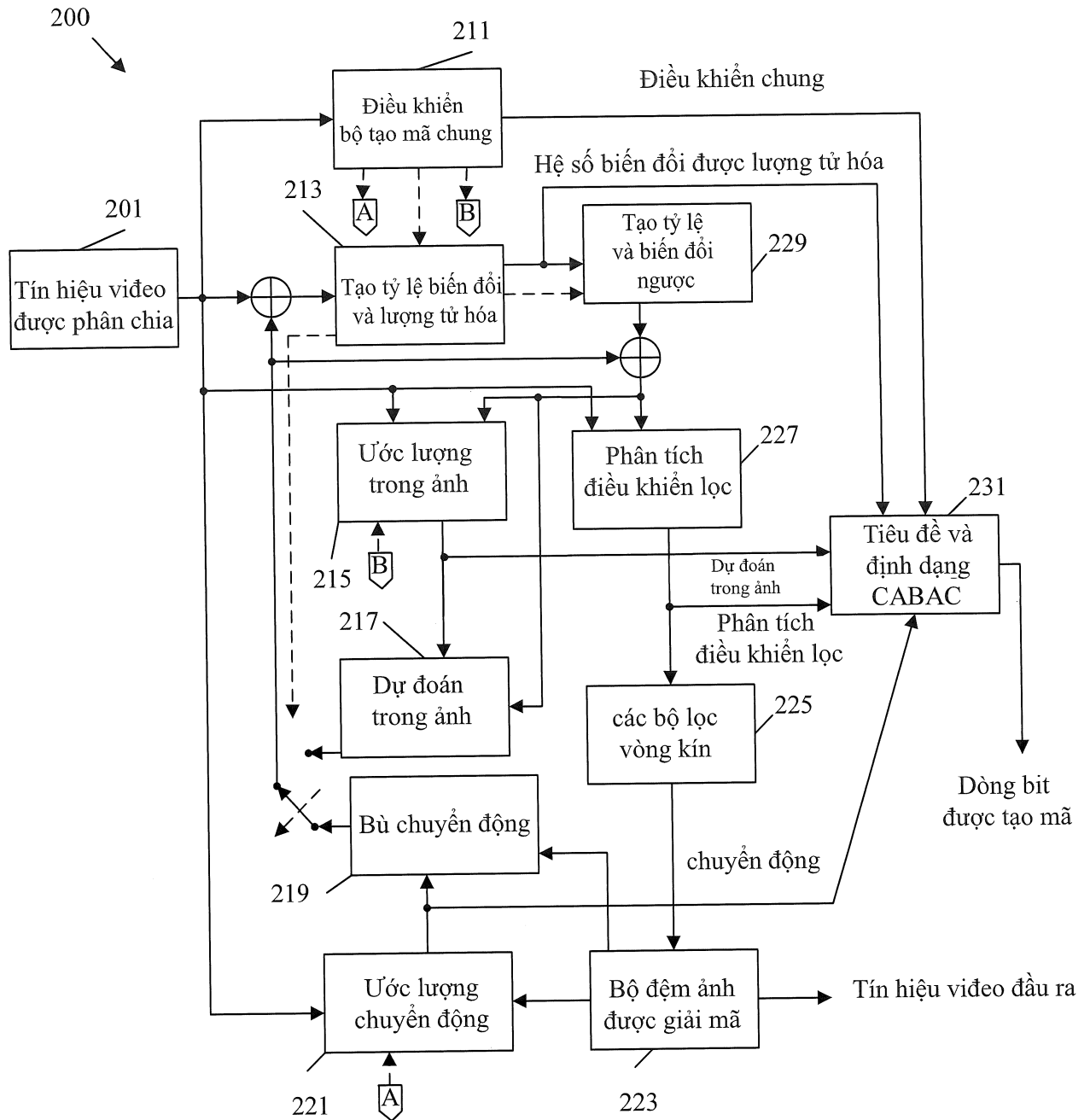


FIG. 2

3/12

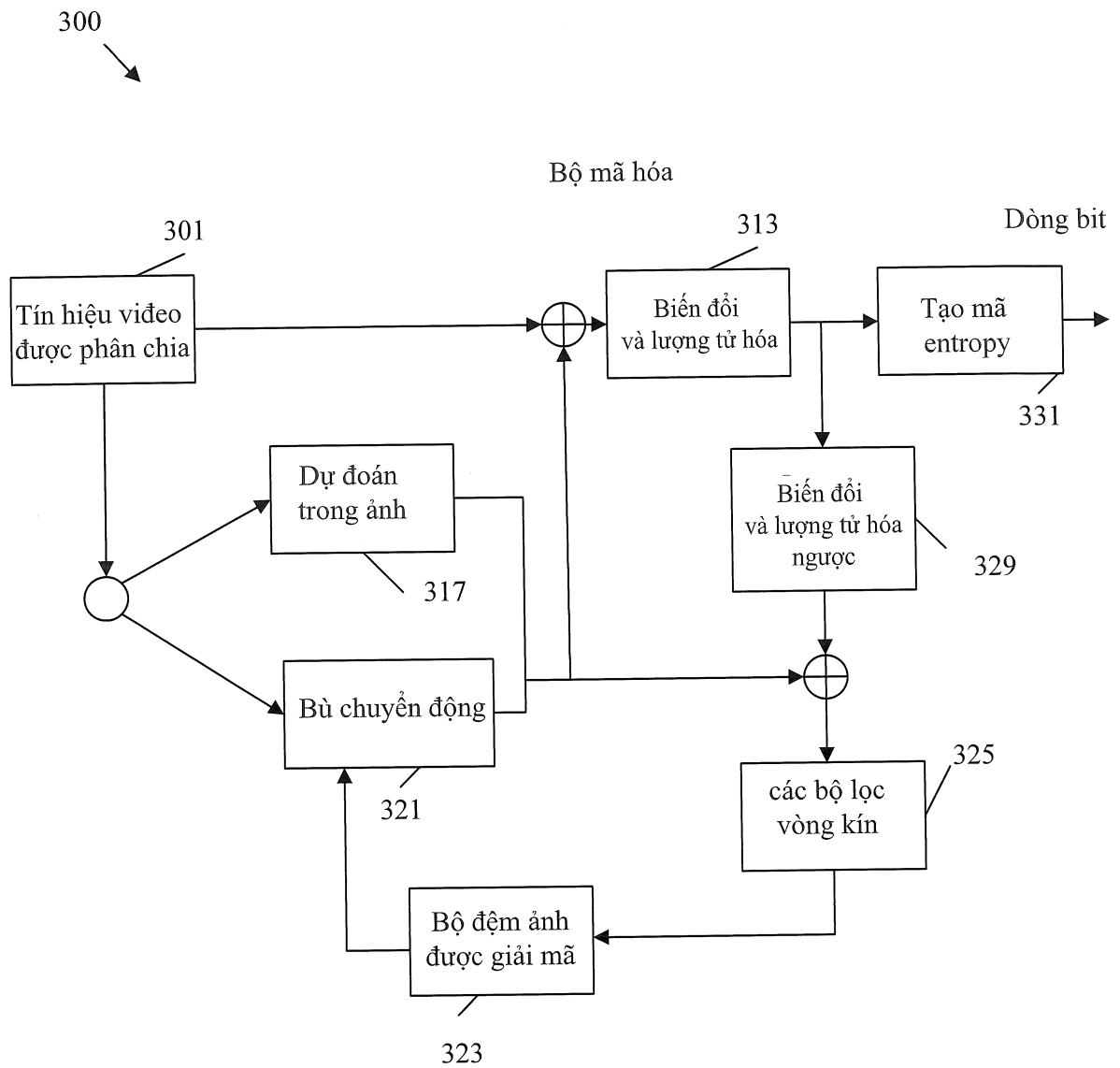


FIG. 3

4/12

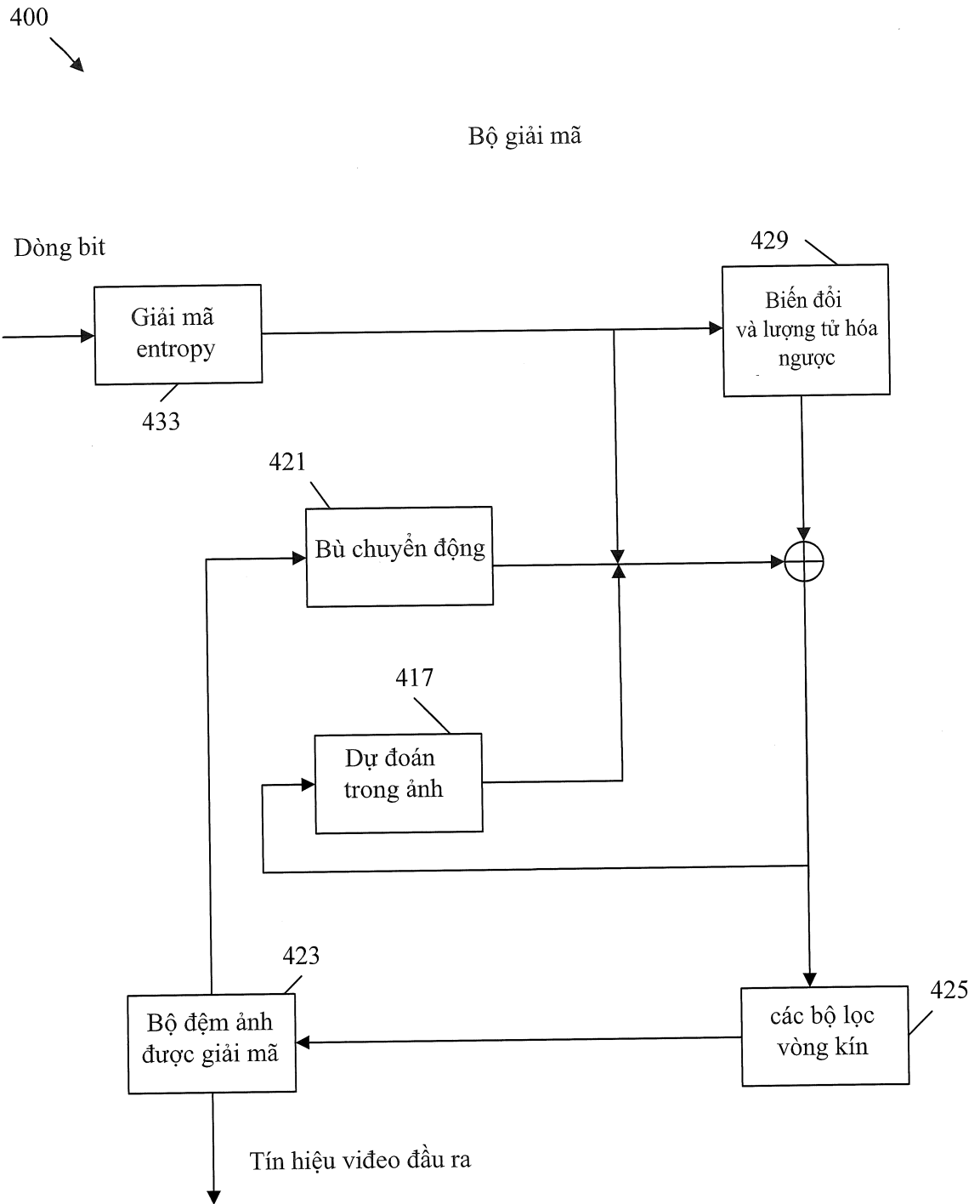


FIG. 4

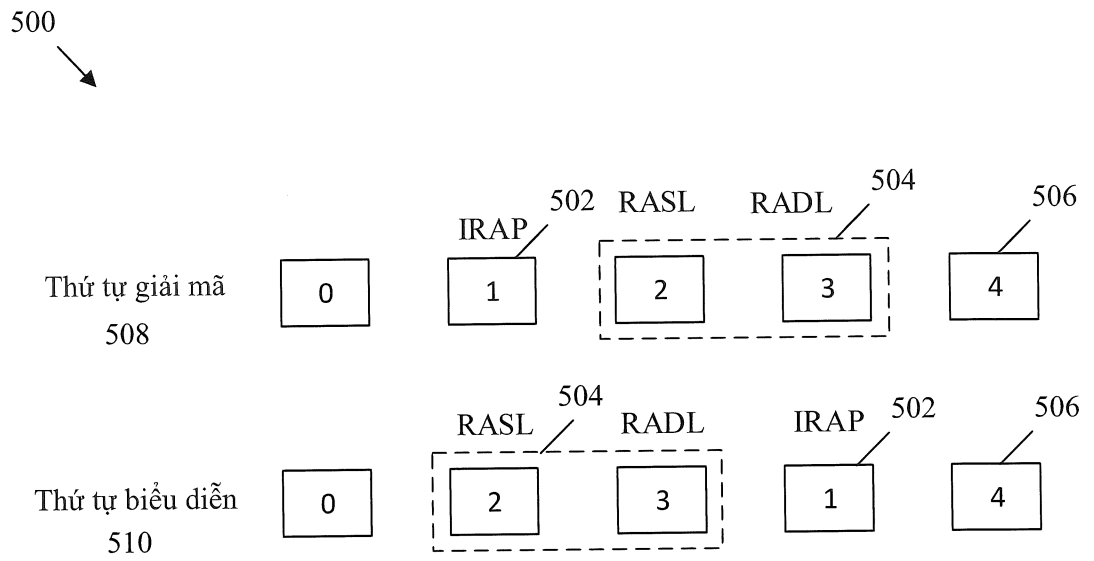


FIG. 5

6/12



FIG. 6 A

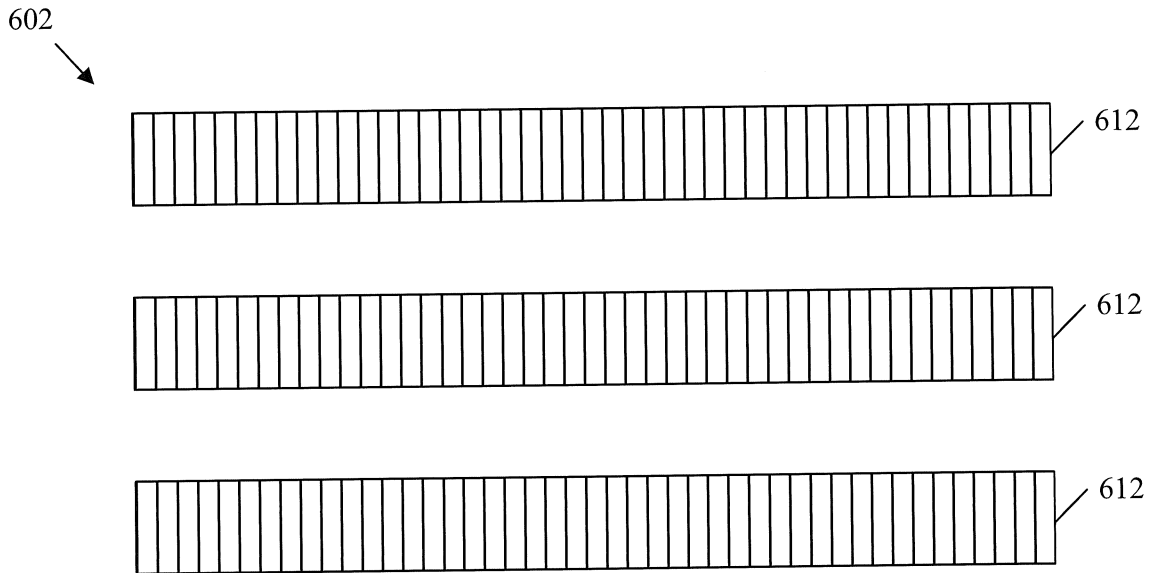


FIG. 6 B

7/12

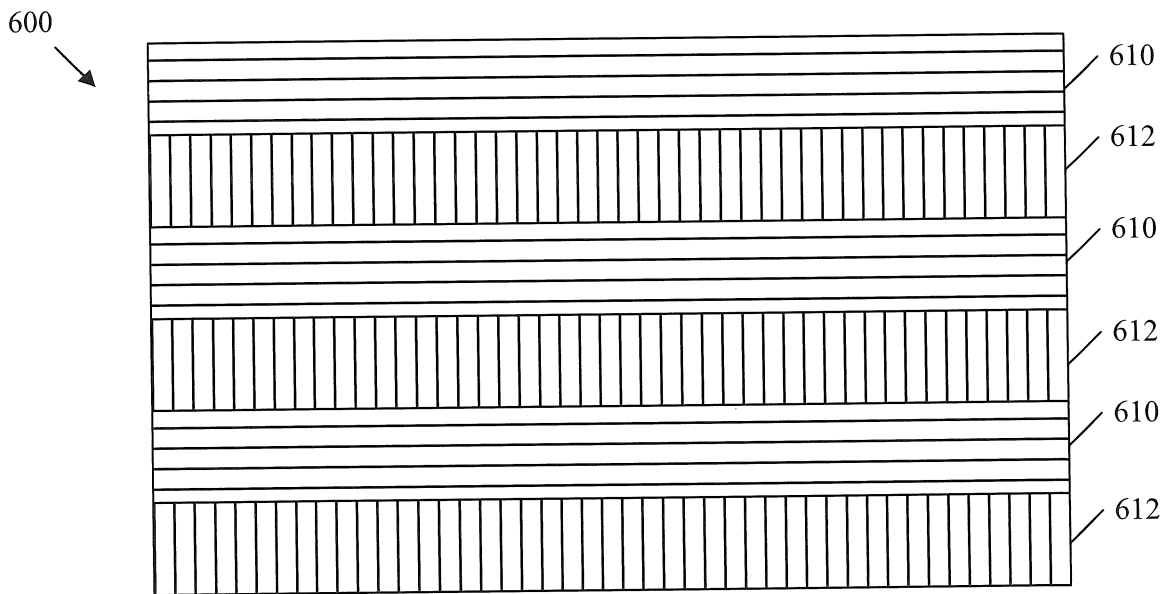


FIG. 6 C

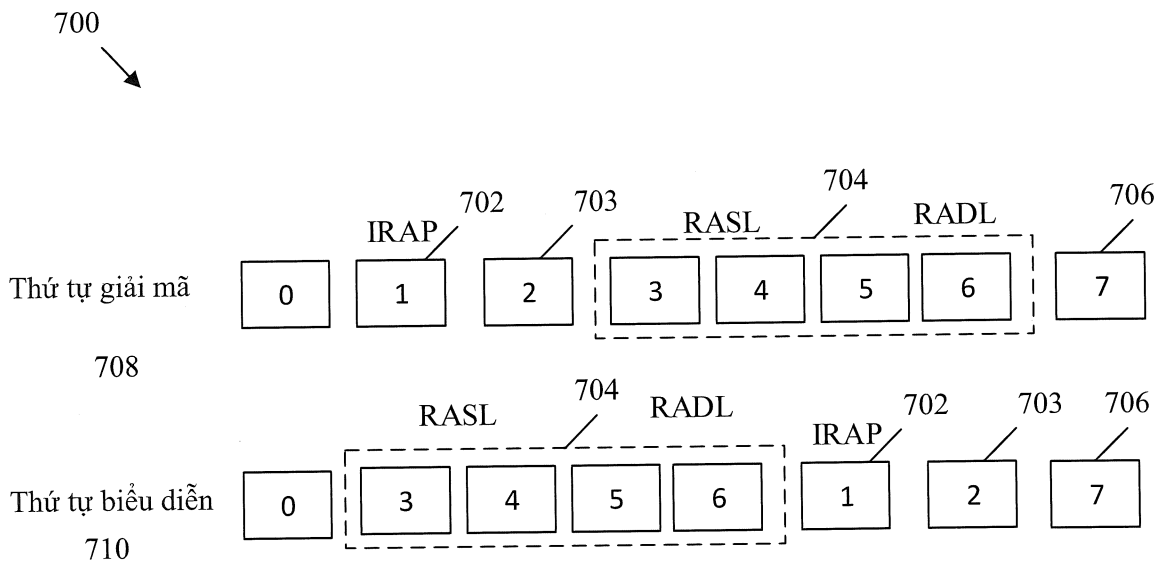


FIG. 7

8/12

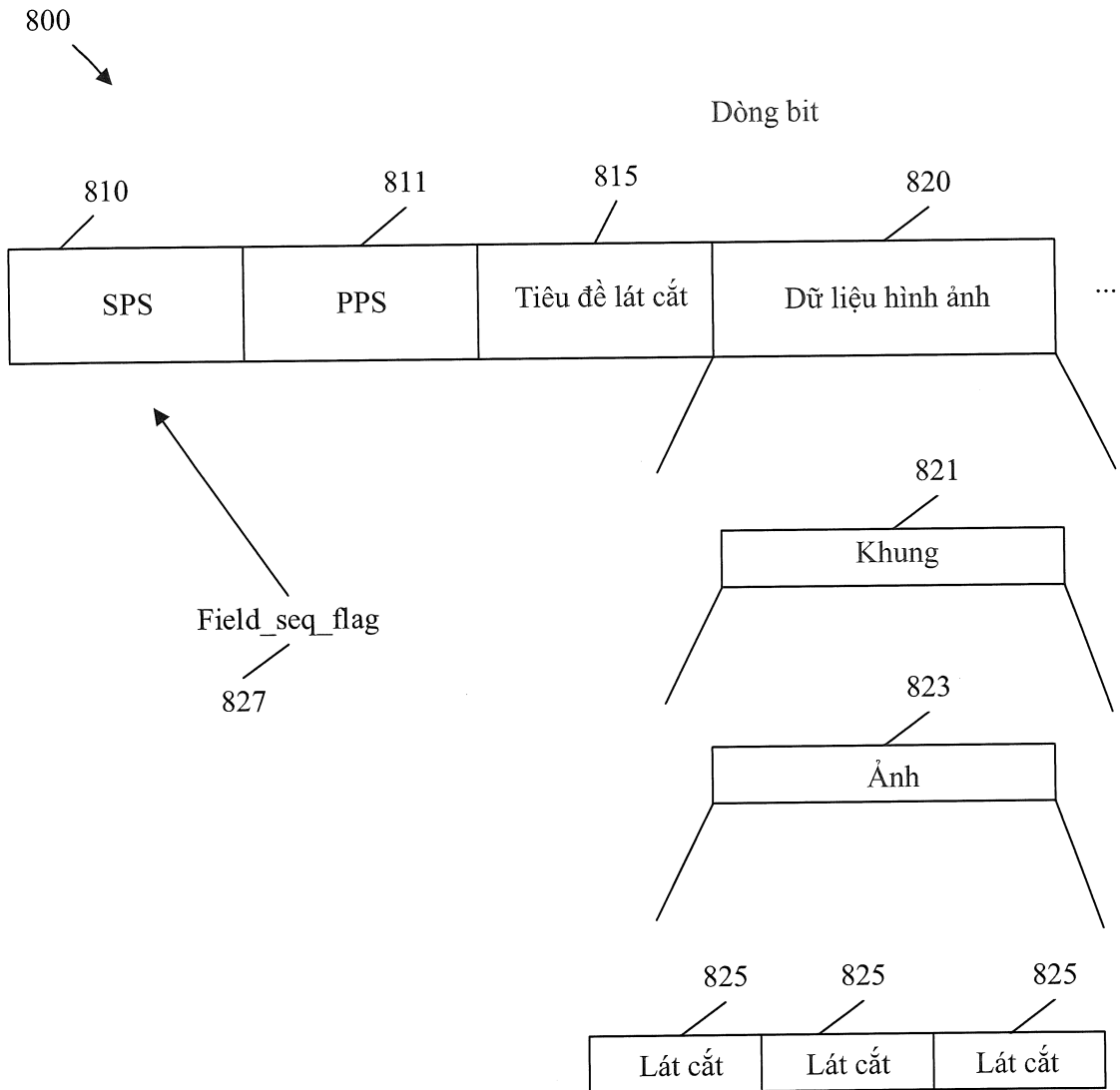


FIG. 8

9/12

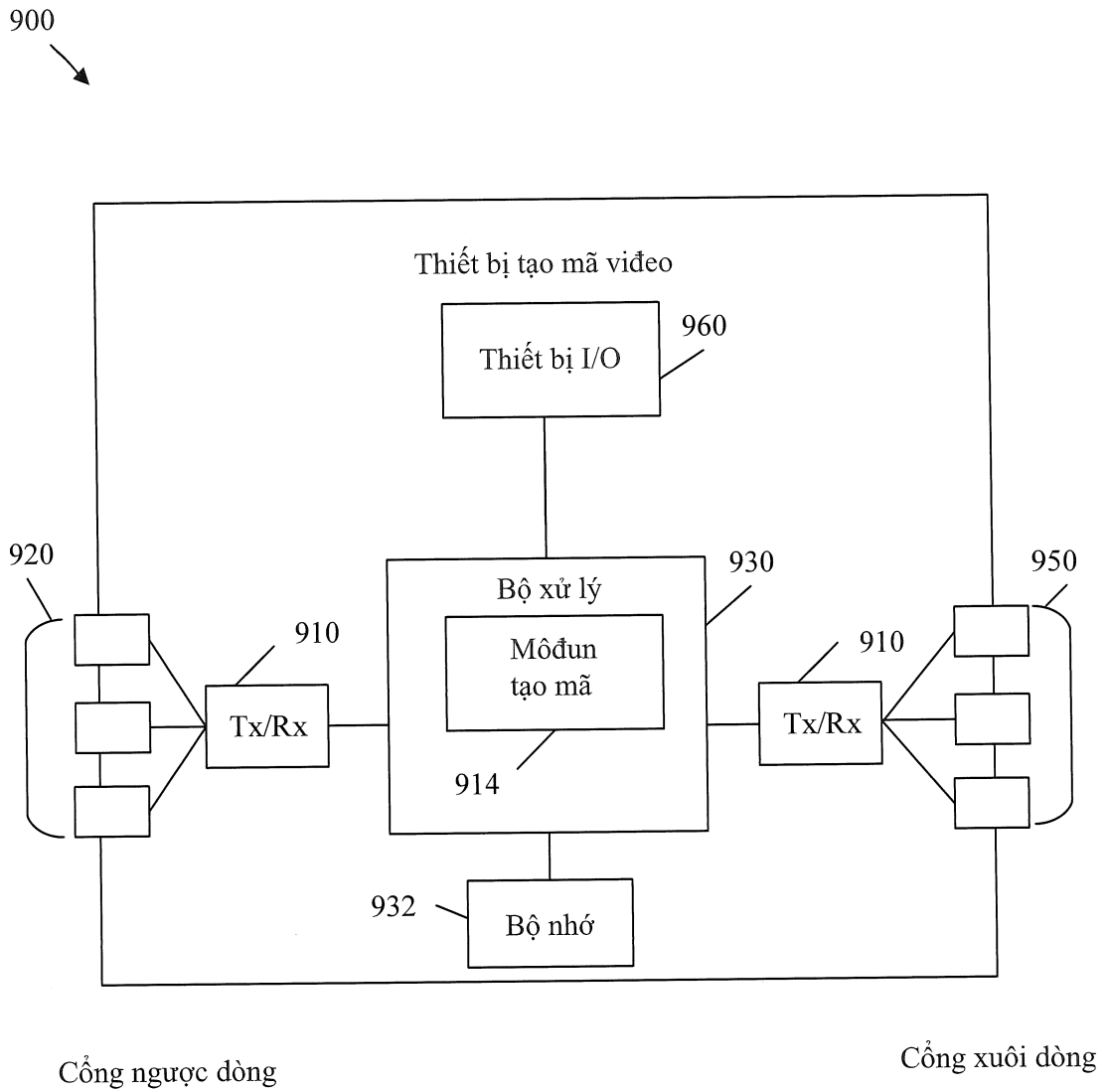


FIG. 9

10/12

1000

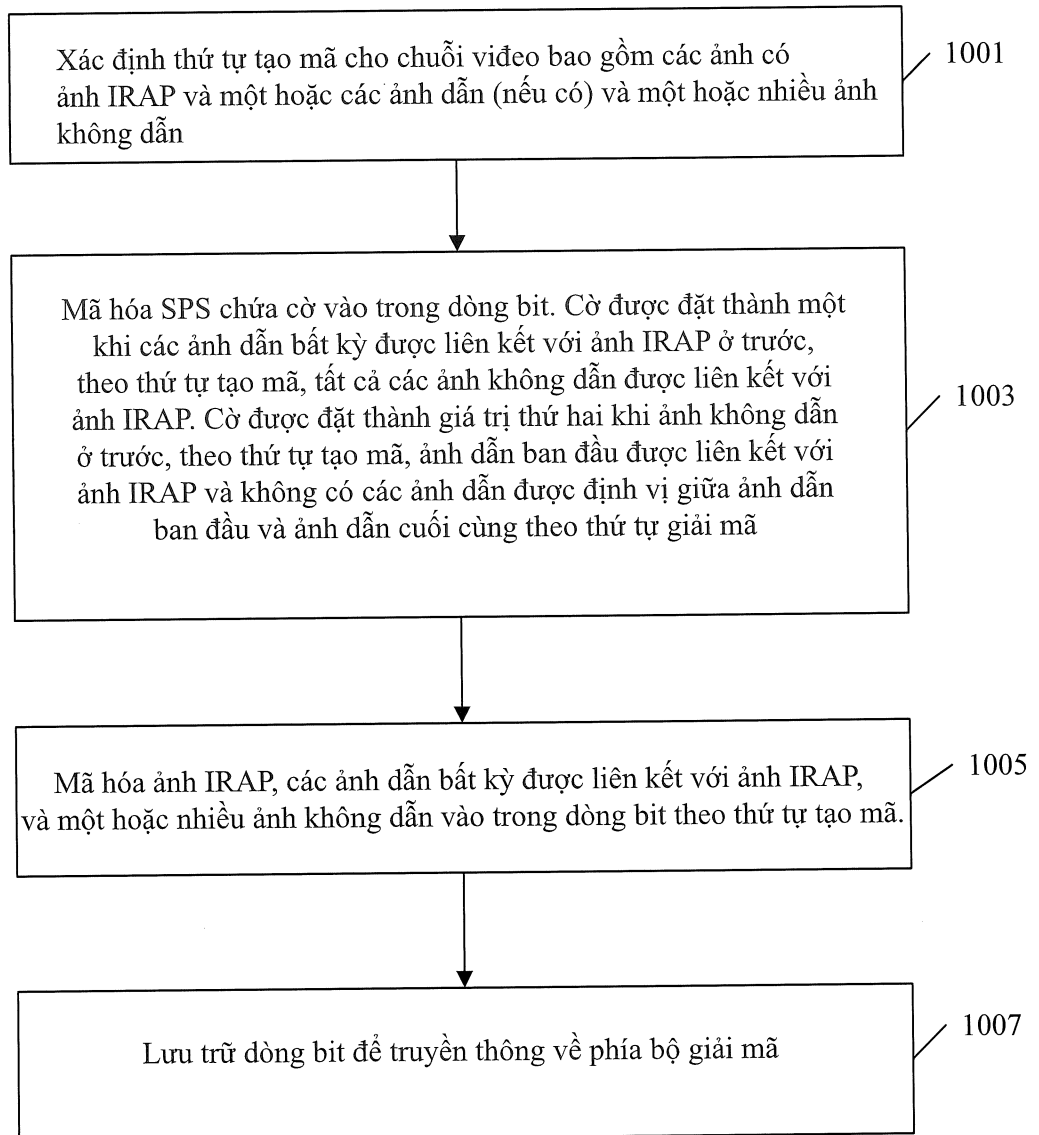


FIG. 10

11/12

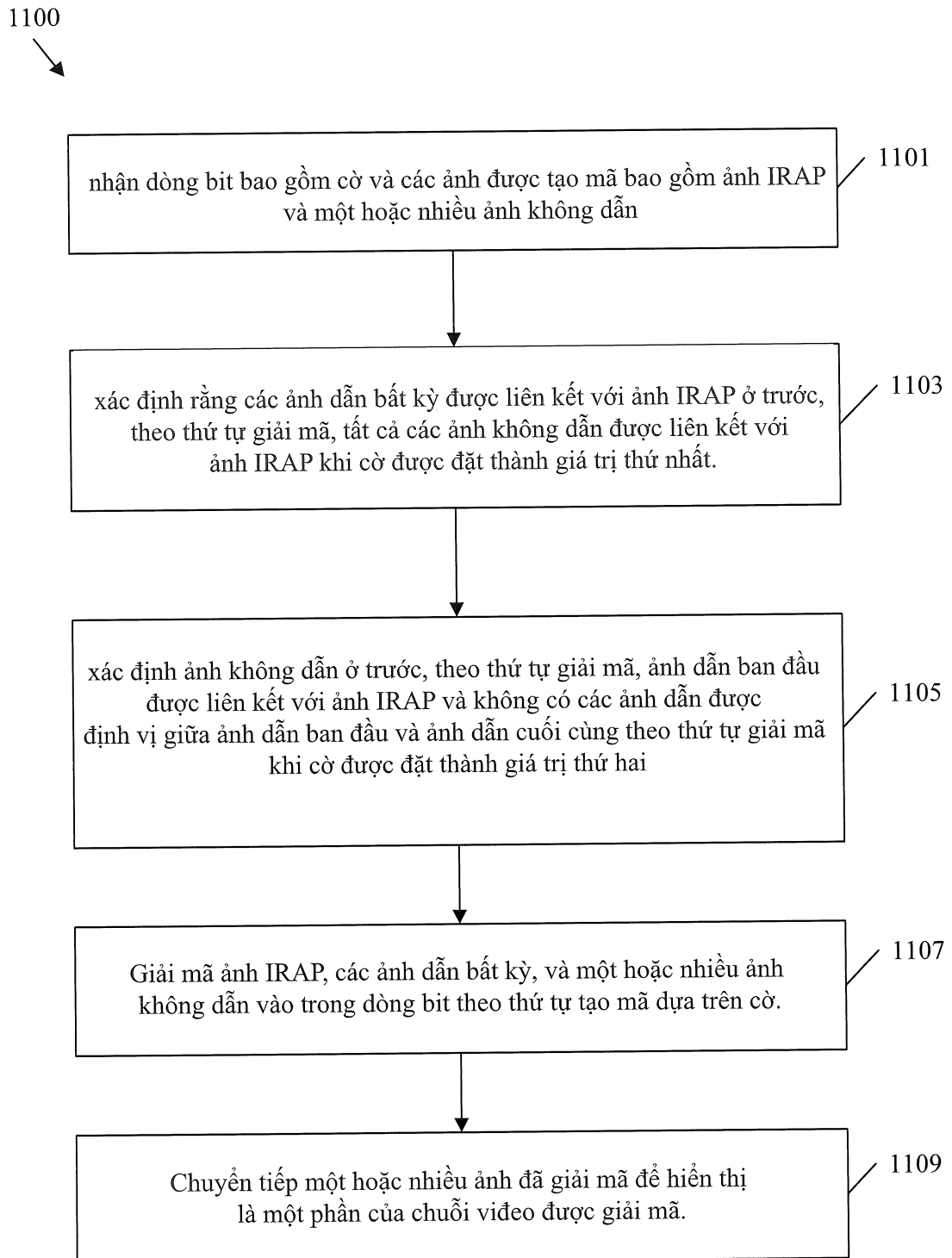


FIG. 11

12/12

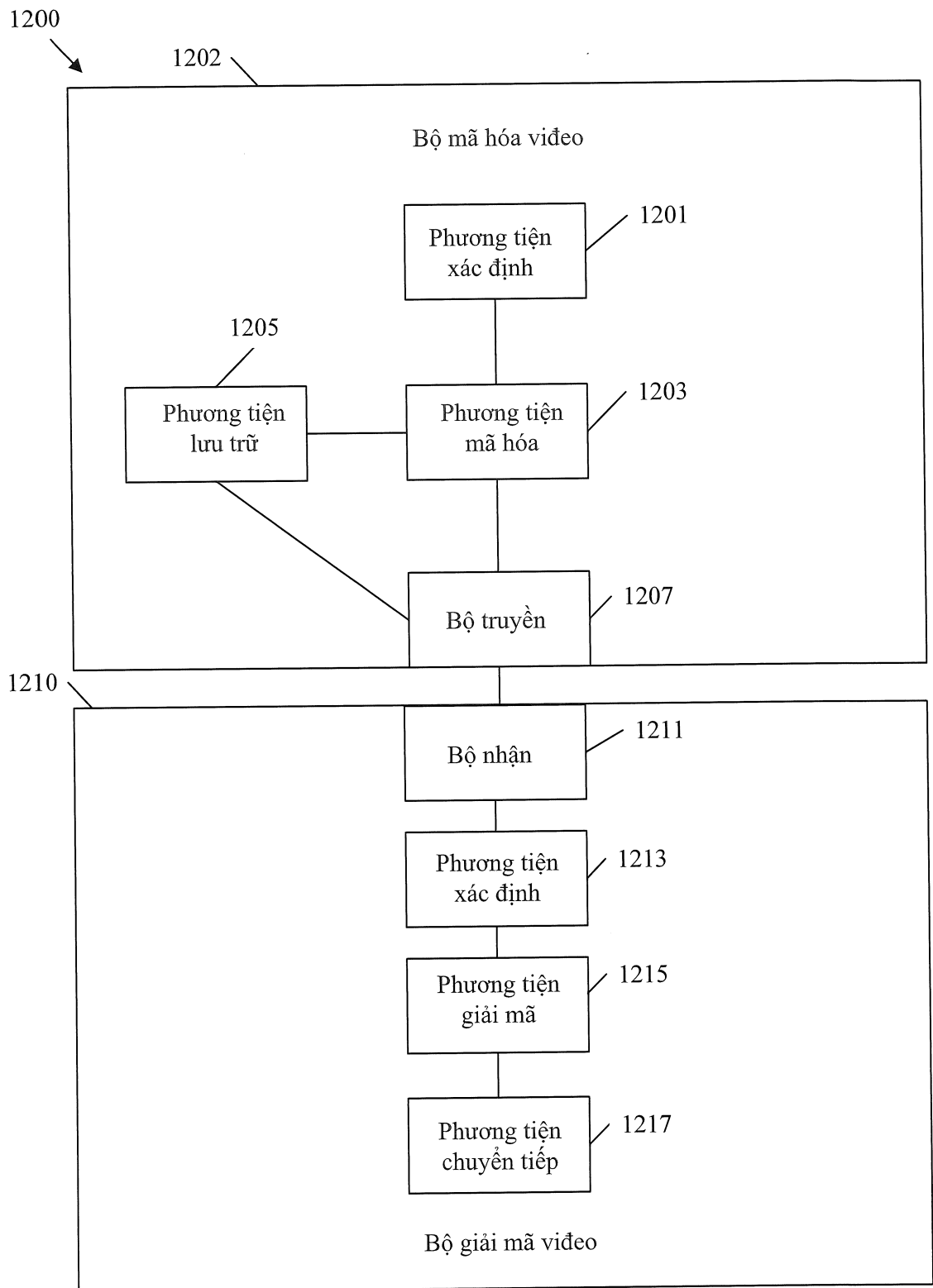


FIG. 12