



- (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
1-0052875
(51)^{2006.01} G06T 9/00; H04N 7/46; H04N 7/32; (13) B
H04N 1/64; H04N 7/12
-

- (21) 1-2022-02460 (22) 05/08/2020
(86) PCT/US2020/045037 05/08/2020 (87) WO 2021/061281 01/04/2021
(30) 62/905,150 24/09/2019 US
(45) 27/10/2025 451 (43) 25/07/2022 412A
(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)
Huawei Administration Building Bantian, Longgang District, Shenzhen, Guangdong
518129, P.R. China
(72) HENDRY, Fnu (ID); WANG, Ye-Kui (US); CHEN, Jianle (CN).
(74) Công ty TNHH một thành viên Sở hữu trí tuệ VCCI (VCCI-IP CO.,LTD)
-

- (54) PHƯƠNG PHÁP GIẢI MÃ, PHƯƠNG PHÁP MÃ HÓA, THIẾT BỊ GIẢI MÃ,
THIẾT BỊ MÃ HÓA, THIẾT BỊ MÃ, HỆ THỐNG, PHƯƠNG TIỆN MÃ, PHƯƠNG
TIỆN LƯU TRỮ ĐỌC ĐƯỢC BỞI MÁY TÍNH VÀ BỘ MÃ

(21) 1-2022-02460

(57) Sáng chế đề cập đến phương pháp giải mã, phương pháp mã hóa, thiết bị giải mã, thiết bị mã hóa, thiết bị mã, hệ thống mã, phương tiện mã, phương tiện lưu trữ đọc được bằng máy tính, phương tiện lưu trữ và bộ mã. Phương pháp giải mã bao gồm thu dòng bit video bao gồm đơn vị ảnh (PU - picture unit), PU bao gồm đơn vị lớp trừu tượng mạng (NAL - network abstraction layer) thông tin tiêu đề ảnh (PH - picture header) mà có kiểu đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL - video coding layer) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID -identifier) lớp của đơn vị PH NAL bằng ID lớp của đơn vị NAL lớp mã hóa video (VCL) được liên kết và ID theo thời gian của đơn vị PH NAL bằng ID theo thời gian của PU; và giải mã ảnh được mã hóa từ PU để thu nhận ảnh được giải mã.

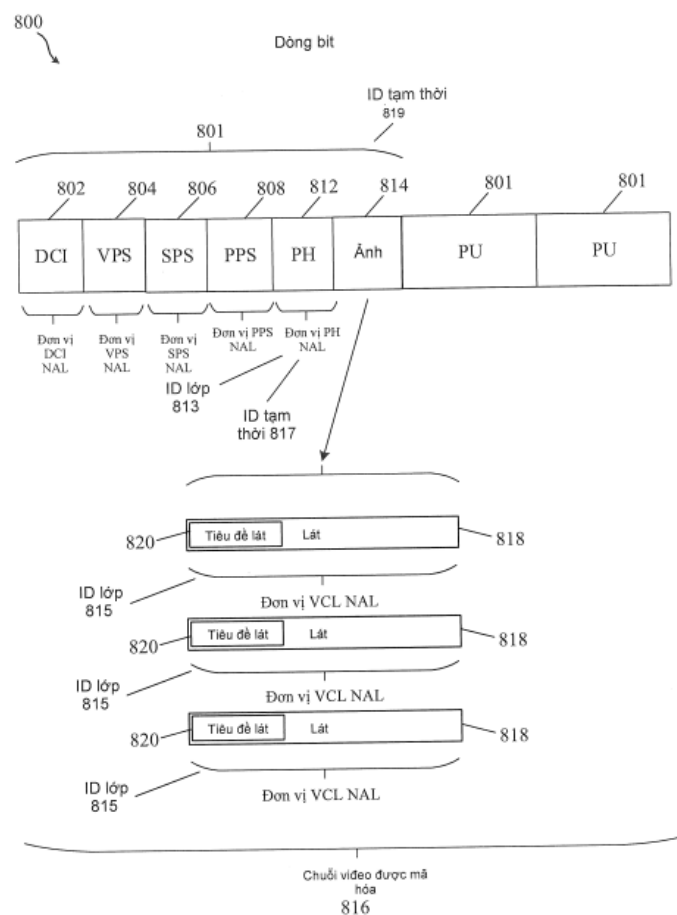


FIG. 8

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các kỹ thuật để xác định điểm bắt đầu của ảnh trong dòng bit trong mã hóa video. Cụ thể hơn, sáng chế đảm bảo rằng điểm bắt đầu của ảnh có thể được xác định chính xác trong dòng bit đa lớp trong mã hóa video.

Tình trạng kỹ thuật của sáng chế

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

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất đề xuất phương pháp giải mã được thực hiện bởi bộ giải mã video, bao gồm: thu, bởi bộ giải mã video, dòng bit video bao gồm đơn vị ảnh (PU - picture unit), PU bao gồm đơn vị lớp trừu tượng mạng (NAL - network abstraction layer) thông tin tiêu đề ảnh (PH - picture header) mà có kiểu đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL - video coding layer) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID - identifier) lớp của đơn vị PH NAL bằng ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng ID theo thời gian của PU; và giải mã, bởi bộ giải mã video, ảnh được mã hóa từ PU để thu nhận ảnh được giải mã.

Phương pháp cung cấp các kỹ thuật mà đảm bảo đơn vị ảnh (PU) bao gồm chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trừu tượng mạng (NAL) PH có kiểu đơn vị PH NAL (PH_NUT) bằng với ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng với ID theo thời gian của PU bao gồm đơn vị PH NAL. Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

Một cách tùy chọn, theo bất kỳ trong số các khía cạnh nêu trên, cách thức thực hiện của khía cạnh đề xuất rằng PU bao gồm chỉ một thông tin tiêu đề ảnh (PH), và trong đó ảnh được mã hóa chỉ là ảnh trong PU.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng việc nhận dạng lớp của ảnh được mã hóa dựa trên ID lớp của đơn vị PH NAL, và nhận dạng vị trí ở thời điểm của ảnh được mã hóa trong dòng bit video dựa trên ID theo thời gian của đơn vị PH NAL.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng đơn vị VCL NAL được liên kết bao gồm lát của ảnh được mã hóa.

Một cách tùy chọn, theo bất kỳ trong số các khía cạnh nêu trên, cách thức thực hiện của khía cạnh đề xuất rằng đơn vị NAL thông tin tiêu đề ảnh là đơn vị NAL không lớp lớp mã hóa video.

Một cách tùy chọn, theo bất kỳ trong số các khía cạnh nêu trên, cách thức thực hiện của khía cạnh đề xuất rằng đơn vị NAL thông tin tiêu đề ảnh bao gồm chỉ một PH.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng ID lớp của đơn vị PH NAL được chỉ định là `nuh_layer_id`.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng ID lớp của đơn vị VCL NAL nhận dạng lớp mà bao gồm đơn vị VCL NAL được liên kết.

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

Khía cạnh thứ hai đề xuất phương pháp mã hóa được thực hiện bởi bộ mã hóa video, phương pháp này bao gồm: tạo ra, bởi bộ mã hóa video, đơn vị ảnh (PU) bao gồm đơn vị lớp trừu tượng mạng (NAL) thông tin tiêu đề ảnh (PH) mà có kiểu đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID) lớp của đơn vị PH NAL bằng ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng ID theo thời gian của PU; mã hóa, bởi bộ mã hóa video, PU thành dòng bit video; và lưu trữ, bởi bộ mã hóa video, dòng bit video cho việc truyền thông tới bộ giải mã video.

Phương pháp cung cấp các kỹ thuật mà đảm bảo đơn vị ảnh (PU) bao gồm chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trừu tượng mạng (NAL) PH có kiểu đơn vị PH NAL (PH_NUT) bằng với ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng với ID theo thời gian của PU bao gồm đơn vị PH NAL. Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

Một cách tùy chọn, theo bất kỳ trong số các khía cạnh nêu trên, cách thức thực hiện của khía cạnh đề xuất rằng PU bao gồm chỉ một thông tin tiêu đề ảnh (PH-picture header) và ảnh được mã hóa, và trong đó ảnh được mã hóa chỉ là ảnh trong PU.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng ID lớp của đơn vị PH NAL nhận dạng lớp của ảnh được mã hóa trong dòng bit video, và trong đó ID theo thời gian của đơn vị PH NAL nhận dạng vị trí ở thời điểm của ảnh được mã hóa trong dòng bit đa lớp.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng đơn vị VCL NAL được liên kết bao gồm lát của ảnh được mã hóa, và trong đó đơn vị PH NAL là đơn vị NAL không phải VCL.

Khía cạnh thứ ba đề cập đến thiết bị giải mã, thiết bị giải mã bao gồm: bộ thu được tạo cấu hình để thu dòng bit video bao gồm đơn vị ảnh (PU), PU bao gồm đơn vị lớp trừu tượng mạng (NAL) thông tin tiêu đề ảnh (PH) có kiểu đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID) lớp của đơn vị PH NAL bằng ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng ID theo thời gian của PU; bộ nhớ được ghép nối với bộ thu, bộ nhớ lưu trữ các lệnh máy; và bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực thi các lệnh máy để khiến cho thiết bị giải mã: giải mã ảnh được mã hóa từ PU để thu nhận ảnh được giải mã.

Thiết bị giải mã cung cấp các kỹ thuật mà đảm bảo đơn vị ảnh (PU) bao gồm chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trừu tượng mạng (NAL) PH có kiểu đơn vị PH NAL (PH_NUT) bằng với ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng với ID theo thời gian của PU bao gồm đơn vị PH NAL. Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

Một cách tùy chọn, theo bất kỳ trong số các khía cạnh nêu trên, cách thức thực hiện của khía cạnh đề xuất rằng PU bao gồm chỉ một thông tin tiêu đề ảnh (PH-picture header), và trong đó ảnh được mã hóa chỉ là ảnh trong PU.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng bộ xử lý được tạo cấu hình để thực thi các lệnh máy để khiến thiết bị giải mã nhận dạng lớp của ảnh được mã hóa trong dòng bit video dựa trên ID lớp của đơn vị PH NAL, và nhận dạng vị trí theo thời điểm của ảnh được mã hóa trong dòng bit video dựa trên ID theo thời gian của đơn vị PH NAL.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng đơn vị VCL NAL được liên kết bao gồm lát của ảnh được mã hóa.

Khía cạnh thứ tư đề cập đến thiết bị mã hóa. Thiết bị mã hóa bao gồm bộ nhớ chứa các lệnh máy; bộ xử lý được ghép nối với bộ nhớ, bộ xử lý được tạo cấu hình để thực hiện các lệnh máy để khiến thiết bị mã hóa: tạo ra đơn vị ảnh (PU) bao

gồm đơn vị lớp trù tượng mạng (NAL) thông tin tiêu đề ảnh (PH) mà có kiểu đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID) lớp của đơn vị PH NAL bằng ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng ID theo thời gian của PU; mã hóa PU thành dòng bit video; và bộ truyền được ghép nối với bộ xử lý, bộ truyền được tạo cấu hình để truyền dòng bit video tới bộ giải mã video.

Thiết bị mã hóa cung cấp các kỹ thuật mà đảm bảo đơn vị ảnh (PU) bao gồm chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trù tượng mạng (NAL) PH có kiểu đơn vị PH NAL (PH_NUT) bằng với ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng với ID theo thời gian của PU bao gồm đơn vị PH NAL. Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

Một cách tùy chọn, theo bất kỳ trong số các khía cạnh nêu trên, cách thức thực hiện của khía cạnh đề xuất rằng PU bao gồm chỉ một thông tin tiêu đề ảnh (PH-picture header) và ảnh được mã hóa, và trong đó ảnh được mã hóa chỉ là ảnh trong PU.

Một cách tùy chọn, trong khía cạnh bất kỳ trước đó, một cách thực hiện khác của khía cạnh đề xuất rằng ID lớp của đơn vị PH NAL nhận dạng lớp của ảnh được mã hóa trong dòng bit video, và trong đó ID theo thời gian của đơn vị PH NAL nhận dạng vị trí theo thời điểm của ảnh được mã hóa trong dòng bit video.

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

Thiết bị mã cung cấp các kỹ thuật mà đảm bảo đơn vị ảnh (PU) bao gồm chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trù tượng mạng (NAL) PH có kiểu đơn vị PH NAL

(PH_NUT) bằng với ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng với ID theo thời gian của PU bao gồm đơn vị PH NAL. Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

Một cách tùy chọn, theo bất kỳ trong số các khía cạnh nêu trên, cách thức thực hiện của khía cạnh đề xuất màn hình có cấu trúc để hiển thị ảnh được giải mã.

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

Hệ thống cung cấp các kỹ thuật mà đảm bảo đơn vị ảnh (PU) bao gồm chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trừu tượng mạng (NAL) PH có kiểu đơn vị PH NAL (PH_NUT) bằng với ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng với ID theo thời gian của PU bao gồm đơn vị PH NAL. Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

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

Phương tiện mã cung cấp các kỹ thuật mà đảm bảo đơn vị ảnh (PU) bao gồm chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trừu tượng mạng (NAL) PH có kiểu đơn vị PH NAL (PH_NUT) bằng ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị PH NAL bằng với ID theo thời gian của PU bao gồm đơn vị PH NAL.

Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

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 phương án 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 ràng hơn từ phần mô tả chi tiết sau đây kết hợp với các hình vẽ kèm theo và bộ yêu cầu bảo hộ.

Mô tả vắn tắt các hình vẽ

Để hiểu đầy đủ hơn về sáng chế, có thể tham chiếu tới 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 tham chiếu giống nhau ký hiệu các bộ phận giống nhau.

FIG.1 là lưu đồ về phương pháp ví dụ để mã hóa tín hiệu video.

FIG.2 là sơ đồ giản lược của hệ số mã hóa và giải mã (codec) ví dụ để mã video.

FIG.3 là sơ đồ giản lược minh họa ví dụ bộ mã hóa video.

FIG.4 là sơ đồ giản lược minh họa ví dụ bộ giải mã video.

FIG.5 minh họa ví dụ về mã hóa đa lớp đối với khả năng biến đổi tỷ lệ không gian.

FIG.6 minh họa dòng bit video có cấu trúc để thực hiện kỹ thuật làm mới giải mã từng bậc (GDR - gradual decoding refresh).

FIG.7 là sơ đồ minh họa tìm kiếm chuyển động không mong muốn khi sử dụng hạn chế của bộ mã hóa để hỗ trợ GDR.

FIG.8 minh họa phương án của dòng bit video.

FIG.9 là phương án về phương pháp giải mã dòng bit video được mã hóa.

FIG.10 là phương án về phương pháp mã hóa dòng bit video được mã hóa.

FIG.11 là sơ đồ giản lược của thiết bị mã hóa video.

FIG.12 là sơ đồ giản lược của phương án về phương tiện mã.

Mô tả chi tiết sáng chế

Cần hiểu ngay từ đầu rằng mặc dù cách thức thực hiện minh họa của một hoặc nhiều phương án được mô tả dưới đây, nhưng các hệ thống và/hoặc phương pháp được bộc lộ có thể được thực hiện bằng cách sử dụng bất kỳ kỹ thuật nào hiện đã biết hay đang tồn tại. Sáng chế không bị giới hạn ở cách thức thực hiện minh họa, các hình vẽ, và các kỹ thuật được minh họa dưới đây, bao gồm các thiết kế và cách thức thực hiện ví dụ được minh họa và được mô tả ở đây, nhưng có thể được cải biến trong phạm vi của yêu cầu bảo hộ đi kèm cùng với phạm vi đầy đủ của các phần tương đương của chúng.

Các thuật ngữ sau đây được định nghĩa như sau trừ khi được sử dụng trong ngữ cảnh khác ở đây. Cụ thể, các định nghĩa sau đây nhằm mục đích làm rõ ràng thêm sáng chế. Tuy nhiên, các thuật ngữ này có thể được mô tả khác nhau trong các ngữ cảnh khác nhau. Theo đó, các thuật ngữ sau đây cần được xem xét là sự bổ sung và không cần được xét đến việc làm giới hạn bất kỳ định nghĩa nào khác của các phần mô tả được đưa ra cho các thuật ngữ này.

Dòng bit là chuỗi của các bit bao gồm dữ liệu video mà được nén cho việc truyền giữa bộ mã hóa và bộ giải mã. Bộ mã hóa là thiết bị mà được tạo cấu hình để sử dụng các quá trình mã hóa để nén dữ liệu video thành dòng bit. Bộ giải mã là thiết bị mà được tạo cấu hình để sử dụng các quá trình giải mã để khôi phục dữ liệu video từ dòng bit cho việc hiển thị. Ảnh là mảng các mẫu độ chói và/hoặc mảng các mẫu sắc độ mà tạo ra khung hoặc trường của nó. Ảnh mà được mã hóa hoặc được giải mã có thể được coi là ảnh hiện tại để thảo luận rõ ràng. Ảnh tham chiếu là ảnh mã chứa các mẫu tham chiếu mà có thể được sử dụng khi mã hóa các ảnh khác bằng cách tham chiếu theo dự đoán liên đới và/hoặc dự đoán lớp liên đới. Danh sách ảnh tham chiếu là danh sách của các ảnh tham chiếu được sử dụng cho việc dự đoán liên đới và/hoặc dự đoán lớp liên đới. Một vài hệ thống mã hóa video sử dụng hai danh sách ảnh tham chiếu, mà có thể được ký hiệu là danh sách ảnh tham chiếu 1 và danh sách ảnh tham chiếu 0. Cấu trúc danh sách ảnh tham chiếu là cấu trúc cú pháp khả lập địa chỉ mà chứa nhiều danh sách ảnh tham chiếu. Dự đoán liên đới là cơ chế mã hóa các mẫu của ảnh hiện tại bằng cách tham chiếu đến các mẫu được chỉ báo trong ảnh tham chiếu khác với ảnh hiện tại trong đó hình ảnh tham chiếu và hình ảnh hiện tại nằm trong cùng lớp. Mục cấu trúc danh sách ảnh tham chiếu là vị trí khả lập địa chỉ trong cấu trúc danh sách ảnh tham chiếu mà chỉ báo ảnh tham chiếu được kết hợp với danh sách ảnh tham chiếu. Thông tin tiêu đề

lát là một phần của lát được mã hóa mà chứa các phần tử dữ liệu liên quan đến tất cả dữ liệu video trong ô được biểu diễn trong lát. Tập hợp tham số ảnh (PPS-picture parameter set) là tập hợp tham số mà chứa dữ liệu liên quan đến toàn bộ ảnh. Cụ thể hơn, PPS là cấu trúc cú pháp mà chứa các phần tử cú pháp mà áp dụng tới 0 hoặc nhiều hơn toàn bộ các ảnh được mã hóa như được xác định bởi phần tử cú pháp được tìm thấy trong mỗi thông tin tiêu đề ảnh. Tập hợp tham số chuỗi (SPS-sequence parameter set) là tập hợp tham số mà chứa dữ liệu liên quan đến chuỗi của cá ảnh. Đơn vị truy nhập (AU-access unit) là tập hợp của một hoặc nhiều ảnh được mã hóa được kết hợp với cùng thời gian hiển thị (ví dụ: cùng số đếm thứ tự ảnh) để xuất ra từ bộ đệm ảnh được giải mã (DPB-decoded picture buffer) (ví dụ: để hiển thị tới người dùng). Dấu phân cách đơn vị truy nhập (AUD-access unit delimiter) là chỉ báo hoặc cấu trúc dữ liệu được sử dụng để chỉ báo sự bắt đầu của AU hoặc biên giữa các AU. Chuỗi video được giải mã là chuỗi của các ảnh mà đã được khôi phục bởi bộ giải mã để chuẩn bị hiển thị tới người dùng.

Các từ viết tắt sau được sử dụng ở đây, khối cây mã hóa (CTB-Coding Tree Block), đơn vị cây mã hóa (CTU-Coding Tree Unit), đơn vị mã hóa (CU-Coding Unit), chuỗi video được mã hóa (CVS-Coded Video Sequence), nhóm chuyên gia video liên kết (JVET-Joint Video Experts Team), tập hợp ô có điều kiện chuyển động (MCTS-Motion-Constrained Tile Set), đơn vị truyền lớn nhất (MTU-Maximum Transfer Unit), lớp trừu tượng mạng (NAL-Network Abstraction Layer), số đếm thứ tự ảnh (POC-Picture Order Count), tập hợp tham số ảnh (PPS-Picture Parameter Set), tải trọng chuỗi Byte thô (RBSP-Raw Byte Sequence Payload), tập hợp tham số chuỗi (SPS-Sequence Parameter Set), mã hóa video đa năng (VVC-Versatile Video Coding) và bản thảo hoạt động (WD-Working Draft).

FIG.1 là lưu đồ về phương pháp hoạt động ví dụ 100 để mã hóa tín hiệu video. Cụ thể, tín hiệu video được mã hóa tại bộ mã hóa. Quá trình mã hóa nén tín hiệu video bằng cách sử dụng các cơ chế khác nhau để làm giảm kích cỡ tệp tin video. Kích cỡ tệp tin nhỏ hơn cho phép tệp tin video được nén cần được truyền tới phía người dùng, trong khi đó làm giảm thông tin tiêu đề bằng thông được liên kết. Bộ giải mã sau đó giải mã tệp tin video được nén để khôi phục tín hiệu video gốc để hiển thị cho người dùng cuối. Quá trình giải mã thường phản ánh quá trình mã hóa để cho phép bộ giải mã khôi phục một cách nhất quán tín hiệu video.

Ở bước 101, tín hiệu video được nhập vào bộ mã hóa. Ví dụ, tín hiệu video có thể là tệp tin video không được nén được lưu trữ trong bộ nhớ. Theo một

ví dụ khác, tệp tin video có thể được chụp bởi thiết bị chụp hình video, như là camera video, và được mã hóa để hỗ trợ việc phát trực tiếp của video. Tệp tin video có thể bao gồm cả thành phần audio và thành phần video. Thành phần video chứa một chuỗi các khung ảnh mà, khi được xem trong chuỗi, đưa ra ấn tượng trực quan của chuyển động. Các khung chứa các điểm ảnh mà được thể hiện trong các thuật ngữ về ánh sáng, được đề cập đến ở đây là các thành phần độ chói (hoặc các mẫu độ chói), và màu sắc, mà được gọi là các thành phần sắc độ (hoặc các mẫu màu). Trong một số ví dụ, các khung có thể cũng chứa các giá trị độ sâu để hỗ trợ quan sát ba chiều.

Ở bước 103, video được phân vùng thành các khối. Việc phân vùng bao gồm chia các điểm ảnh trong mỗi khung thành các khối vuông và/hoặc các khối chữ nhật cho việc nén. Ví dụ, trong Mã hóa video hiệu quả cao (HEVC) (cũng được biết đến là Phần 2 của H.265 và MPEG-H) khung trước tiên có thể được chia thành các đơn vị cây mã hóa (CTUs), mà là các khối của kích cỡ được xác định trước (ví dụ như, 64 nhân 64 điểm ảnh). CTUs chứa cả các mẫu độ chói và các mẫu sắc độ. Các cây mã hóa có thể được sử dụng để chia CTUs thành các khối và sau đó chia nhỏ để quy các khối cho đến khi các cấu hình đạt được mà hỗ trợ mã hóa nhiều hơn. Ví dụ, các thành phần độ chói của khung có thể được chia nhỏ đến khi các khối riêng biệt chứa các giá trị chiếu sáng tương đối đồng nhất. Hơn nữa, các thành phần sắc độ của khung có thể được chia nhỏ đến khi các khối riêng biệt chứa các giá trị màu tương đối đồng nhất. Theo đó, các cơ chế phân vùng thay đổi phụ thuộc vào nội dung của các khung video.

Ở bước 105, các cơ chế nén khác nhau được sử dụng để nén các khối ảnh được phân vùng ở bước 103. Ví dụ, dự đoán liên đới và/hoặc dự đoán nội bộ có thể được sử dụng. Dự đoán liên đới được thiết kế để tận dụng ưu điểm của thực tế rằng các đối tượng trong cảnh chung có xu hướng xuất hiện trong các khung liên tiếp. Do đó, khối mà thể hiện đối tượng trong khung tham chiếu không cần phải được mô tả lặp lại trong các khung liên kế. Cụ thể, đối tượng, như là Bảng, có thể duy trì trong vị trí không đổi qua nhiều khung. Do đó Bảng được mô tả một lần và các khung liên kế có thể viện dẫn ngược lại đến khung tham chiếu. Các cơ chế so khớp kiểu mẫu có thể được sử dụng để so khớp các đối tượng qua nhiều khung. Hơn nữa, việc di chuyển các đối tượng có thể được thể hiện qua nhiều khung, ví dụ do sự di chuyển đối tượng hoặc di chuyển camera. Đối với ví dụ cụ thể, video có thể thể hiện xe ô tô mà di chuyển qua màn hình qua nhiều khung. Các vectơ chuyển động

có thể được sử dụng để mô tả sự di chuyển này. Vector chuyển động là vector hai chiều mà cung cấp độ dịch từ các tọa độ của đối tượng trong khung cho các tọa độ của đối tượng trong khung tham chiếu. Như vậy, dự đoán liên đới có thể mã hóa khối ảnh trong khung hiện tại dưới dạng tập hợp các vector chuyển động mà chỉ báo độ dịch từ khối tương ứng trong khung tham chiếu.

Các khối mã hóa dự đoán nội bộ trong khung chung. Dự đoán nội bộ có ưu điểm là các thành phần độ chói và sắc độ có xu hướng tập hợp trong khung. Ví dụ, phần vá của màu xanh trong phần của cây có xu hướng được đặt vị trí liên kề với các phần vá tương tự của màu xanh. Dự đoán nội bộ sử dụng nhiều chế độ dự đoán định hướng (ví dụ như, 33 trong HEVC), chế độ phẳng, và chế độ dòng một chiều (DC). Các chế độ định hướng chỉ báo rằng khối hiện tại là tương tự/giống với các mẫu của khối lân cận theo chiều tương ứng. Chế độ phẳng chỉ báo rằng một chuỗi các khối theo hàng/cột (ví dụ như, mặt phẳng) có thể được nội suy dựa trên các khối lân cận tại các cạnh của hàng. Chế độ phẳng, có hiệu lực, chỉ báo việc chuyển đổi suôn sẻ của ánh sáng/màu sắc qua hàng/cột bằng cách sử dụng độ dốc không đổi tương đối trong khi thay đổi các giá trị. Chế độ DC được sử dụng cho việc làm mượt biên và chỉ báo rằng khối là tương tự/giống với giá trị trung bình được kết hợp với các mẫu của tất cả các khối lân cận được kết hợp với các chiều góc của các chế độ dự đoán định hướng. Theo đó, các khối dự đoán nội bộ có thể thể hiện các khối ảnh dưới dạng các giá trị chế độ dự đoán liên quan khác nhau thay vì các giá trị thực. Hơn nữa, các khối dự đoán liên đới có thể thể hiện các khối ảnh dưới dạng các giá trị vector chuyển động thay vì các giá trị thực. Trong trường hợp khác, các khối dự đoán có thể không thể hiện chính xác các khối ảnh trong một số trường hợp. Các độ chênh lệch bất kỳ được lưu trữ trong các khối dư. Các biến đổi có thể được áp dụng cho các khối dư để nén hơn nữa tệp tin.

Ở bước 107, các kỹ thuật lọc khác nhau có thể được áp dụng. Trong HEVC, các bộ lọc được áp dụng theo cơ chế lọc trong vòng. Dự đoán dựa trên khối được đề cập ở trên có thể có kết quả là sự tạo ra của các ảnh khối tại bộ giải mã. Hơn nữa, cơ chế dự đoán dựa trên khối có thể mã hóa khối và sau đó khôi phục khối được mã hóa để sử dụng sau dưới dạng khối tham chiếu. Cơ chế lọc trong vòng áp dụng theo cách lặp các bộ lọc khử tạp âm, các bộ lọc giải khối, các bộ lọc vòng thích ứng, và các bộ lọc độ dịch thích ứng mẫu (SAO - sample adaptive offset) cho các khối/khung. Các bộ lọc này làm giảm bớt các tạo tác ngăn cản để tệp tin được mã hóa có thể được khôi phục một cách chính xác. Hơn nữa, các bộ lọc này

làm giảm bớt các tạo tác trong các khối tham chiếu được khôi phục để các tạo tác ít có khả năng tạo ra các tạo tác bổ sung trong các khối kế tiếp mà được mã hóa dựa trên các khối tham chiếu được khôi phục.

Một khi tín hiệu video được phân vùng, được nén, và được lọc, kết quả là dữ liệu được mã hóa trong dòng bit ở bước 109. Dòng bit bao gồm dữ liệu được đề cập ở trên cũng như báo hiệu dữ liệu bất kỳ mong muốn hỗ trợ sự khôi phục tín hiệu video chính xác ở bộ giải mã. Ví dụ, dữ liệu này có thể bao gồm dữ liệu phân vùng, dữ liệu dự đoán, các khối dư, và các cờ khác nhau mà cung cấp các lệnh mã hóa cho bộ giải mã. Dòng bit có thể được lưu trữ trong bộ nhớ cho việc truyền tới bộ giải mã khi có yêu cầu. Dòng bit có thể cũng phát rộng và/hoặc phát đa hướng tới các bộ giải mã. Việc tạo ra dòng bit là quá trình lặp. Theo đó, các bước 101, 103, 105, 107, và 109 có thể xuất hiện tiếp diễn và/hoặc đồng thời qua nhiều khung và khối. Lệnh được thể hiện trên FIG.1 được thể hiện để làm rõ và thuận tiện cho việc thảo luận, và không nhằm giới hạn ở quá trình mã hóa video cho thứ tự cụ thể.

Bộ giải mã thu dòng bit và bắt đầu quá trình giải mã ở bước 111. Cụ thể, bộ giải mã sử dụng cơ chế giải mã entropy để chuyển đổi dòng bit thành cú pháp và dữ liệu video tương ứng. Bộ giải mã sử dụng dữ liệu cú pháp từ dòng bit để xác định các phân vùng dùng cho các khung ở bước 111. Việc phân vùng cần so khớp với các kết quả của khối mà phân vùng ở bước 103. Mã hóa/giải mã entropy như được sử dụng trong bước 111 được mô tả ở đây. Bộ mã hóa tạo ra nhiều lựa chọn trong quá trình nén, như là lựa chọn các cơ chế phân vùng khối từ các lựa chọn khả thi dựa trên việc định vị theo không gian của các giá trị trong (các) ảnh đầu vào. Việc báo hiệu các lựa chọn chính xác có thể sử dụng một số lượng lớn các ký tự nhị phân. Như được sử dụng ở đây, ký tự nhị phân là giá trị nhị phân mà được xử lý như một biến (ví dụ như, giá trị bit mà có thể thay đổi phụ thuộc vào ngữ cảnh). Việc mã hóa entropy cho phép bộ mã hóa loại bỏ các tùy chọn bất kỳ mà rõ ràng là không khả thi cho trường hợp cụ thể, bỏ đi tập hợp các tùy chọn có thể được cho phép. Mỗi tùy chọn được cho phép sau đó được gán từ mã. Độ dài của các từ mã dựa trên số lượng tùy chọn được cho phép (ví dụ như, một ký tự nhị phân dùng cho hai tùy chọn, hai ký tự nhị phân dùng cho ba đến bốn tùy chọn, v.v.). Bộ mã hóa sau đó mã hóa từ mã dùng cho tùy chọn được lựa chọn. Cơ chế này làm giảm kích cỡ của các từ mã khi các từ mã lớn như mong muốn để chỉ báo duy nhất sự lựa chọn từ tập con nhỏ của các tùy chọn được phép trái ngược với để chỉ báo duy nhất sự lựa chọn từ tập tất cả các tùy chọn khả thi lớn tiềm năng. Bộ giải mã sau đó giải

mã sự lựa chọn bằng cách xác định tập hợp các tùy chọn được phép theo cách thức tương tự cho bộ mã hóa. Bằng cách xác định tập hợp các tùy chọn được phép, bộ giải mã có thể đọc từ mã và xác định sự lựa chọn được thực hiện bởi bộ mã hóa.

Ở bước 113, bộ giải mã thực hiện giải mã khối. Cụ thể, bộ giải mã sử dụng 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 khôi phục các khối ảnh theo sự phân vùng. Các khối dự đoán có thể bao gồm cả các khối dự đoán nội bộ và các khối dự đoán liên đới như được tạo ra ở bộ mã hóa ở bước 105. Các khối ảnh được khôi phục sau đó được đặt trong các khung của báo hiệu video được khôi phục theo dữ liệu phân vùng được xác định ở bước 111. Cú pháp dùng cho bước 113 có thể cũng được báo hiệu trong dòng bit qua mã hóa entropy như được đề cập ở trên.

Ở bước 115, việc lọc được thực hiện trên các khung của tín hiệu video được khôi phục theo cách thức tương tự như bước 107 ở bộ mã hóa. Ví dụ, các bộ lọc khử tạp âm, các bộ lọc khử khối, các bộ lọc vòng thích ứng, và các bộ lọc SAO có thể được áp dụng cho các khung để loại bỏ các tạo tác khối. Một khi các khung được lọc, tín hiệu video có thể được xuất ra màn hình ở bước 117 để quan sát bởi người dùng.

FIG.2 là sơ đồ giản lược của ví dụ hệ thống mã hóa và giải mã (codec) 200 để mã hóa video. Cụ thể, hệ thống codec 200 cung cấp chức năng để hỗ trợ cách thực hiện của phương pháp thao tác 100. Hệ thống codec 200 được tạo ra để mô tả các thành phần được sử dụng trong cả hai bộ mã hóa và bộ giải mã. Hệ thống codec 200 thu và phân vùng tín hiệu video như được đề cập đến đối với các bước 101 và 103 trong phương pháp thao tác 100, mà dẫn đến tín hiệu video được phân vùng 201. Hệ thống codec 200 sau đó nén tín hiệu video được phân vùng 201 thành dòng bit được mã hóa khi đóng vai trò là bộ mã hóa như được đề cập đối với các bước 105, 107, và 109 trong phương pháp 100. Khi đóng vai trò là bộ giải mã, hệ thống codec 200 tạo ra tín hiệu video đầu ra từ dòng bit như được đề cập đối với các bước 111, 113, 115, và 117 trong phương pháp thao tác 100. Hệ thống codec 200 bao gồm thành phần điều khiển bộ mã hóa chung 211, thành phần biến đổi tỷ lệ và lượng tử hóa 213, thành phần ước lượng ảnh nội bộ 215, thành phần dự đoán ảnh nội bộ 217, thành phần bù chuyển động 219, thành phần ước lượng chuyển động 221, thành phần biến đổi tỷ lệ và biến đổi ngược 229, thành phần phân tích điều khiển bộ lọc 227, thành phần lọc trong vòng 225, thành phần bộ đệm ảnh được giải mã 223, và thành phần mã hóa số học nhị phân tích ứng ngữ cảnh và khuôn

dạng tiêu đề (CABAC) 231. Các thành phần này được ghép nối như được thể hiện. Trên FIG.2, các đường màu đen chỉ báo sự di chuyển của dữ liệu cần được mã hóa/giải mã trong khi đó các đường nét gạch chỉ báo sự di chuyển của dữ liệu điều khiển mà điều khiển thao tác của các thành phần khác. Các thành phần của hệ thống codec 200 có thể đều có trong bộ mã hóa. Bộ giải mã có thể bao gồm tập con của các thành phần của hệ thống codec 200. Ví dụ, bộ giải mã có thể bao gồm thành phần dự đoán ảnh nội bộ 217, thành phần bù chuyển động 219, thành phần biến đổi tỷ lệ và biến đổi ngược 229, thành phần lọc trong vòng 225, và thành phần đệm ảnh được giải mã 223. Các thành phần này được mô tả ở đây.

Tín hiệu video được phân vùng 201 là chuỗi video được ghi hình mà được phân vùng thành các khối của các điểm ảnh bởi cây mã hóa. Cây mã hóa sử dụng các chế độ phân chia khác nhau để chia nhỏ khối của các điểm ảnh thành các khối nhỏ hơn của các điểm ảnh. Các khối này có thể sau đó được chia nhỏ tiếp thành các khối nhỏ hơn. Các khối này có thể được gọi là các nút trên cây mã hóa. Các nút họ lớn hơn được phân chia thành các nút con nhỏ hơn. Số lần một nút được chia nhỏ được gọi là độ sâu của nút/cây mã hóa. Các khối được chia có thể nằm trong các đơn vị mã hóa (CUs) trong một số trường hợp. Ví dụ, CU có thể là phần con của CTU mà chứa khối độ chói, (các) khối sắc độ chênh lệch màu đỏ (Cr), và (các) khối sắc độ chênh lệch màu xanh (Cb) theo các lệnh cú pháp tương ứng cho CU. Các chế độ phân chia có thể bao gồm cây nhị phân (BT), cây tam phân (TT), và cây tứ phân (QT) được sử dụng để phân vùng nút thành hai, ba, hoặc bốn nút con, theo cách tương ứng, có các hình dạng khác nhau phụ thuộc vào các chế độ phân chia được sử dụng. Tín hiệu video được phân vùng 201 được chuyển tiếp đến thành phần điều khiển bộ mã hóa chung 211, thành phần biến đổi tỷ lệ và lượng tử hóa 213, thành phần ước lượng ảnh nội bộ 215, thành phần phân tích điều khiển bộ lọc 227, và thành phần ước lượng chuyển động 221 dùng cho việc nén.

Thành phần điều khiển bộ mã hóa chung 211 được tạo cấu hình để ra quyết định dành cho việc mã hóa của các ảnh của chuỗi video thành dòng bit theo các ràng buộc ứng dụng. Ví dụ, thành phần điều khiển bộ mã hóa chung 211 quản lý việc tối ưu hóa của tốc độ bit/kích cỡ dòng bit so với chất lượng khôi phục. Các quyết định đó có thể được thực hiện dựa trên tính phù hợp của không gian lưu trữ/băng thông và các yêu cầu độ phân giải ảnh. Thành phần điều khiển bộ mã hóa chung 211 cũng quản lý việc sử dụng bộ đệm về tốc độ truyền để giảm nhẹ các vấn đề chạy dưới mức và quá mức đệm. Để kiểm soát các vấn đề này, thành phần điều

khiển bộ mã hóa chung 211 quản lý việc phân vùng, dự đoán, và lọc bởi các thành phần khác. Ví dụ, thành phần điều khiển bộ mã hóa chung 211 có thể tăng động độ phức tạp nén để làm tăng độ phân giải và làm tăng sử dụng băng thông hoặc làm giảm độ phức tạp nén để làm giảm độ phân giải và sử dụng băng thông. Do đó, thành phần điều khiển bộ mã hóa chung 211 điều khiển các thành phần khác của hệ thống codec 200 để cân bằng chất lượng khôi phục tín hiệu video với tốc độ bit tương xứng. Thành phần điều khiển bộ mã hóa chung 211 tạo ra dữ liệu điều khiển, mà điều khiển thao tác của các thành phần khác. Dữ liệu điều khiển cũng được chuyển tiếp đến thành phần CABAC và lấy khuôn dạng tiêu đề 231 cần được mã hóa trong dòng bit cho các tham số tín hiệu để giải mã tại bộ giải mã.

Tín hiệu video được phân vùng 201 cũng được gửi đến thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 dùng cho dự đoán liên đới. Khung hoặc lát của tín hiệu video được phân vùng 201 có thể được chia thành nhiều khối video. Thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 thực hiện mã hóa dự đoán liên đới của khối video thu được tương ứng với một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu để cung cấp dự đoán thời gian. Hệ thống codec 200 có thể thực hiện nhiều lần mã hóa, ví dụ như, để lựa chọn chế độ mã hóa phù hợp cho mỗi khối của dữ liệu video.

Thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp, nhưng được minh họa một cách riêng biệt nhằm mục đích xác định khái niệm. Ước lượng chuyển động, được thực hiện bởi thành phần ước lượng chuyển động 221, là quá trình tạo ra các vector chuyển động, mà ước lượng chuyển động của các khối video. Vector chuyển động, ví dụ, có thể chỉ báo sự di lệch của đối tượng được mã hóa tương ứng với khối dự đoán. Khối dự đoán là khối mà được tìm thấy khớp gần với khối được mã hóa, tính về mặt độ lệch điểm ảnh. Khối dự đoán có thể cũng được gọi là khối tham chiếu. Độ lệch điểm ảnh này có thể được xác định bởi tổng của độ lệch tuyệt đối (SAD), tổng của bình phương độ lệch (SSD), hoặc các chỉ số độ lệch khác. HEVC sử dụng nhiều đối tượng được mã hóa bao gồm CTU, các khối cây mã hóa (các CTB), và các CU. Ví dụ, CTU có thể được chia thành các CTB, mà sau đó có thể được chia thành các CB để đưa vào CU. CU có thể được mã hóa như là đơn vị dự đoán mà chứa dữ liệu dự đoán và/hoặc đơn vị biến đổi (TU) mà chứa dữ liệu dư được biến đổi đối với CU. Bộ ước lượng chuyển động 221 tạo ra các vector chuyển động, các đơn vị dự đoán, và các TU bằng cách sử dụng phân tích méo tỷ lệ như là một phần của xử lý tối ưu hóa méo tỷ

lệ. Ví dụ, bộ ước lượng chuyển động 221 có thể xác định các khối tham chiếu, các vector chuyển động, v.v đối với khối/khung hiện tại, và có thể lựa chọn các khối tham chiếu, các vector chuyển động, v.v mà có đặc tính méo tỷ lệ tốt nhất. Các đặc tính biến dạng tỷ lệ tốt nhất làm cân bằng cả chất lượng khôi phục video (ví dụ, lượng tổn hao dữ liệu do nén) với hiệu quả mã hóa (ví dụ, kích cỡ của mã hóa cuối cùng).

Trong một số ví dụ, hệ thống codec 200 có thể tính toán các giá trị cho các vị trí điểm ảnh số nguyên phụ của các hình ảnh tham chiếu được lưu trữ trong thành phần bộ đệm ảnh được giải mã 223. Ví dụ, hệ thống codec video 200 có thể nội suy các giá trị của các vị trí điểm ảnh một phần tư, các vị trí điểm ảnh một phần tám, hoặc các vị trí điểm ảnh tỷ lệ khác của hình ảnh tham chiếu. Do đó, thành phần ước lượng chuyển động 221 có thể thực hiện tìm kiếm chuyển động tương ứng với các vị trí điểm ảnh đầy đủ và các vị trí điểm ảnh theo tỷ lệ và xuất ra vector chuyển động với độ chính xác điểm ảnh theo tỷ lệ. Bộ ước lượng chuyển động 221 tính toán vector chuyển động đối với đơn vị dự đoán của khối video trong lát được mã hóa liên đới bằng cách so sánh vị trí của đơn vị dự đoán với vị trí của khối dự đoán của ảnh tham chiếu. Bộ ước lượng chuyển động 221 xuất ra vector chuyển động được tính toán như là dữ liệu chuyển động tới bộ CABAC và tạo khuôn dạng thông tin tiêu đề 231 để mã hóa và chuyển động tới bộ bù chuyển động 219.

Việc bù chuyển động, được thực hiện bởi bộ bù chuyển động 219, có thể liên quan đến việc tìm hoặc tạo ra khối dự đoán dựa trên vector chuyển động được xác định bởi bộ ước lượng chuyển động 221. Một cách lặp lại, thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp theo chức năng, trong một số ví dụ. Sau khi thu vector chuyển động đối với đơn vị chuyển động của khối video hiện tại, bộ bù chuyển động 219 có thể xác định vị trí khối dự đoán mà vector chuyển động trỏ tới. Sau đó, khối video dự được tạo ra bằng cách trừ các giá trị điểm ảnh của khối dự đoán từ các giá trị điểm ảnh của khối video hiện tại đang được mã hóa, tạo thành các giá trị chênh lệch điểm ảnh. Nói chung, bộ ước lượng chuyển động 221 thực hiện việc ước lượng chuyển động so với các thành phần độ chói và bộ bù chuyển động 219 sử dụng các vector chuyển động được tính toán dựa trên các thành phần độ chói cho cả các thành phần sắc độ và thành phần độ chói. Khối dự đoán và khối dự được chuyển tiếp để tạo thành thành phần biến đổi tỷ lệ và lượng tử hóa 213.

Tín hiệu video được phân vùng 201 cũng được gửi đến thành phần ước lượng ảnh nội bộ 215 và thành phần dự đoán ảnh nội bộ 217. Đối với thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219, thành phần ước lượng ảnh nội bộ 215 và thành phần dự đoán ảnh nội bộ 217 có thể được tích hợp cao, nhưng được minh họa riêng biệt cho mục đích lấy định nghĩa. Thành phần ước lượng ảnh nội bộ 215 và thành phần dự đoán ảnh nội bộ 217 dự đoán nội bộ cho khối hiện tại tương ứng với các khối trong khung hiện tại, như sự thay thế cho dự đoán liên đới được thực hiện bởi thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 giữa các khung, như được mô tả ở trên. Cụ thể là, thành phần ước lượng ảnh nội bộ 215 xác định chế độ dự đoán nội bộ để sử dụng để mã hóa khối hiện tại. Trong một số ví dụ, thành phần ước lượng ảnh nội bộ 215 lựa chọn chế độ dự đoán nội bộ phù hợp để mã hóa khối hiện tại từ nhiều chế độ dự đoán nội bộ được thử nghiệm. Các chế độ dự đoán nội bộ được lựa chọn sau đó được chuyển tiếp đến thành phần CABAC và lấy khuôn dạng tiêu đề 231 để mã hóa.

Ví dụ, thành phần ước lượng ảnh nội bộ 215 tính toán các giá trị biến dạng tỷ lệ sử dụng phân tích biến dạng tỷ lệ cho các chế độ dự đoán nội bộ được thử nghiệm khác nhau, và lựa chọn chế độ dự đoán nội bộ mà có các đặc tính biến dạng tỷ lệ tốt nhất trong số các chế độ được thử nghiệm. Phân tích biến dạng tỷ lệ thường xác định lượng biến dạng (hoặc sai số) giữa khối được mã hóa và khối gốc, chưa được mã hóa mà đã được mã hóa để đưa ra khối được mã hóa, cũng như tỷ lệ bit (ví dụ như, số lượng bit) được sử dụng để đưa ra khối được mã hóa. Thành phần ước lượng ảnh nội bộ 215 tính toán các tỷ lệ từ các biến dạng tỷ lệ và các tỷ lệ cho các khối được mã hóa khác nhau để xác định chế độ dự đoán nội bộ nào đưa ra giá trị biến dạng tỷ lệ tốt nhất cho khối. Ngoài ra, thành phần ước lượng ảnh nội bộ 215 có thể được tạo cấu hình để mã hóa các khối độ sâu của ánh xạ độ sâu bằng cách sử dụng chế độ tạo mẫu độ sâu (DMM) dựa trên tối ưu hóa biến dạng tỷ lệ (RDO).

Thành phần dự đoán ảnh nội bộ 217 có thể tạo ra khối dư từ khối dự đoán dựa trên các chế độ dự đoán nội bộ được lựa chọn được xác định bởi thành phần ước lượng ảnh nội bộ 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 sự khác biệt về các giá trị giữa khối dự đoán và khối gốc, được thể hiện dưới dạng ma trận. Khối dư sau đó được chuyển tiếp đến thành phần biến đổi tỷ lệ và lượng tử hóa 213. Thành

phần ước lượng ảnh nội bộ 215 và thành phần dự đoán ảnh nội bộ 217 có thể thao tác trên cả thành phần độ chói và sắc độ.

Thành phần biến đổi tỷ lệ và lượng tử hóa 213 được tạo cấu hình để nén hơn nữa khối dư. Thành phần biến đổi tỷ lệ và lượng tử hóa 213 áp dụng biến đổi, như là biến đổi cosin rời rạc (DCT), biến đổi sin rời rạc (DST), hoặc biến đổi tương tự về mặt khái niệm, cho khối dư, đưa ra khối video bao gồm các giá trị hệ số biến đổi dư. Các phép biến đổi Wavelet, biến đổi nguyên, biến đổi băng con hoặc các loại biến đổi khác cũng có thể được sử dụng. Việc biến đổi có thể chuyển đổi thông tin dư từ miền giá trị điểm ảnh sang miền biến đổi, chẳng hạn như miền tần số. Bộ lượng tử hóa và biến đổi tỷ lệ 213 cũng có cấu trúc để biến đổi tỷ lệ thông tin dư được biến đổi, ví dụ dựa trên tần số. Việc biến đổi tỷ lệ này bao gồm việc áp dụng hệ số biến đổi tỷ lệ tới thông tin dư sao cho thông tin tần số khác nhau được lượng hóa tại các mức độ chi tiết khác nhau, mà có thể ảnh hưởng đến chất lượng trực quan cuối cùng của video được khôi phục. Thành phần biến đổi tỷ lệ và lượng tử hóa 213 cũng được tạo cấu hình để lượng tử hóa các hệ số biến đổi để làm giảm hơn nữa tốc độ bit. Xử lý lượng tử hóa có thể làm giảm độ sâu bit được kết hợp với một số hoặc tất cả các hệ số. Mức độ lượng tử hóa có thể được cải biến bằng cách điều chỉnh tham số lượng tử hóa. Trong một vài ví dụ, bộ lượng tử hóa và biến đổi tỷ lệ 213 có thể sau đó thực hiện việc quét ma trận bao gồm các hệ số biến đổi được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa được chuyển tiếp tới bộ CABAC và tạo khuôn dạng thông tin tiêu đề 231 để được mã hóa trong dòng bit.

Thành phần biến đổi tỷ lệ và biến đổi ngược 229 áp dụng thao tác dành riêng của thành phần biến đổi tỷ lệ và lượng tử hóa 213 để hỗ trợ ước lượng chuyển động. Thành phần biến đổi tỷ lệ và biến đổi ngược 229 áp dụng biến đổi tỷ lệ ngược, biến đổi, và/hoặc lượng tử hóa để khôi phục khối dư trong miền điểm ảnh, ví dụ như, để sử dụng sau như khối tham chiếu mà có thể trở thành khối dự đoán cho một khối hiện tại khác. Thành phần ước lượng chuyển động 221 và/hoặc thành phần bù chuyển động 219 có thể tính toán khối tham chiếu bằng cách thêm khối dư trở lại vào khối dự đoán tương ứng cho việc sử dụng trong ước lượng chuyển động của khối/khung đi sau. Các bộ lọc được áp dụng cho các khối tham chiếu được khôi phục để giảm nhẹ các tạo tác được tạo ra trong khi biến đổi tỷ lệ, lượng tử hóa, và biến đổi. Các tạo tác này ngược lại có thể khiến dự đoán không chính xác (và tạo ra các tạo tác bổ sung) khi các khối kế tiếp được dự đoán.

Thành phần phân tích điều khiển bộ lọc 227 và thành phần bộ lọc trong vòng 225 áp dụng các bộ lọc cho các khối dư và/hoặc cho các khối ảnh được khôi phục. Ví dụ, khối dư được biến đổi từ thành phần biến đổi tỷ lệ và biến đổi ngược 229 có thể được kết hợp với khối dự đoán tương ứng từ thành phần dự đoán ảnh nội bộ 217 và/hoặc thành phần bù chuyển động 219 để khôi phục khối ảnh gốc. Các bộ lọc có thể sau đó được áp dụng cho khối ảnh được khôi phục. Trong một số ví dụ, thay vào đó các bộ lọc có thể được áp dụng cho các khối dư. Đối với các thành phần khác trên FIG.2, thành phần phân tích điều khiển bộ lọc 227 và thành phần bộ lọc trong vòng 225 được tích hợp cao và có thể được thực hiện cùng nhau, nhưng được mô tả riêng biệt nhằm mục đích lấy khái niệm. Các bộ lọc được áp dụng cho các khối tham chiếu được khôi phục được áp dụng cho các vùng không gian cụ thể và bao gồm nhiều tham số để điều chỉnh cách thức các bộ lọc này được áp dụng. Thành phần phân tích điều khiển bộ lọc 227 phân tích các khối tham chiếu được khôi phục để xác định trong đó các bộ lọc này nên được áp dụng và thiết đặt các tham số tương ứng. Dữ liệu này được chuyển tiếp đến thành phần CABAC và lấy khuôn dạng tiêu đề 231 như dữ liệu điều khiển bộ lọc để mã hóa. Thành phần bộ lọc trong vòng 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 khử nhiễu, bộ lọc SAO, và bộ lọc vòng thích ứng. Các bộ lọc này có thể được áp dụng trong miền không gian/điểm ảnh (ví dụ như, trên khối điểm ảnh được khôi phục) hoặc trong miền tần số, phụ thuộc vào ví dụ.

Khi hoạt động như bộ mã hóa, khối ảnh được khôi phục được lọc, khối dư, và/hoặc khối dự đoán được lưu trữ trong thành phần bộ đệm ảnh được giải mã 223 cho việc sử dụng sau trong ước lượng chuyển động như được đề cập ở trên. Khi hoạt động như bộ giải mã, thành phần bộ đệm ảnh được giải mã 223 lưu trữ và chuyển tiếp các khối được khôi phục và được lọc tới màn hình như một phần của tín hiệu video đầu ra. Thành phần bộ đệm ảnh được giải mã 223 có thể là thiết bị ghi nhớ bất kỳ có khả năng lưu trữ các khối dự đoán, các khối dư, và/hoặc các khối ảnh được khôi phục.

Thành phần CABAC và lấy khuôn dạng tiêu đề 231 thu dữ liệu từ các thành phần khác nhau của hệ thống codec 200 và mã hóa dữ liệu này thành dòng bit được mã hóa cho việc truyền tới bộ giải mã. Cụ thể, thành phần CABAC và lấy khuôn dạng tiêu đề 231 tạo ra các tiêu đề khác nhau để mã hóa dữ liệu điều khiển, như là dữ liệu điều khiển chung và dữ liệu điều khiển bộ lọc. Hơn nữa, dữ liệu dự

đoán, bao gồm dự đoán nội bộ và dữ liệu chuyển động, cũng như dữ liệu dư theo dạng dữ liệu hệ số biến đổi được lượng tử hóa đều được mã hóa trong dòng bit. Dòng bit cuối cùng bao gồm tất cả thông tin được mong muốn có bởi bộ giải mã để khôi phục tín hiệu video được phân vùng gốc 201. Thông tin này có thể cũng bao gồm các bảng chỉ số chế độ dự đoán nội bộ (cũng được gọi là từ các bảng ánh xạ mã), các định nghĩa của mã hóa các ngữ cảnh cho các khối khác nhau, các chỉ báo của các chế độ dự đoán nội bộ khả thi nhất, chỉ báo của thông tin phân vùng, v.v. Dữ liệu này có thể được mã hóa bằng cách sử dụng mã hóa entropy. Ví dụ, thông tin có thể được mã hóa bằng cách sử dụng mã hóa độ dài biến thích ứng ngữ cảnh (CAVLC), CABAC, mã hóa số học nhị phân thích ứng ngữ cảnh dựa trên cú pháp (SBAC), mã hóa entropy phân vùng khoảng khả thi (PIPE), hoặc một kỹ thuật mã hóa entropy khác. Tuân theo việc mã hóa entropy, dòng bit được mã hóa có thể được truyền đến một thiết bị khác (ví dụ như, bộ giải mã video 28) hoặc được lưu trữ cho việc truyền hoặc tìm kiếm sau.

FIG.3 là sơ đồ khối minh họa ví dụ bộ mã hóa video 300. Bộ mã hóa video 300 có thể được sử dụng để thực hiện các chức năng mã hóa của hệ thống codec 200 và/hoặc thực hiện các bước 101, 103, 105, 107, và/hoặc 109 của phương pháp thao tác 100. Bộ mã hóa 300 phân vùng tín hiệu video đầu vào, dẫn đến tín hiệu video được phân vùng 301, mà về cơ bản tương tự như tín hiệu video được phân vùng 201. Tín hiệu video được phân vùng 301 sau đó được nén và được mã hóa thành dòng bit bởi các thành phần của bộ mã hóa 300.

Cụ thể, tín hiệu video được phân vùng 301 được chuyển tiếp đến thành phần dự đoán ảnh nội bộ 317 dùng cho dự đoán nội bộ. Thành phần dự đoán ảnh nội bộ 317 có thể về cơ bản tương tự như thành phần ước lượng ảnh nội bộ 215 và thành phần dự đoán ảnh nội bộ 217. Tín hiệu video được phân vùng 301 cũng được chuyển tiếp đến thành phần bù chuyển động 321 cho dự đoán liên đới dựa trên các khối tham chiếu trong thành phần bộ đệm ảnh được giải mã 323. Thành phần bù chuyển động 321 có thể về cơ bản tương tự như thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219. Các khối dự đoán và các khối dư từ thành phần dự đoán ảnh nội bộ 317 và thành phần bù chuyển động 321 được chuyển tiếp đến thành phần biến đổi và lượng tử hóa 313 dùng cho việc biến đổi và lượng tử hóa của các khối dư. Thành phần biến đổi và lượng tử hóa 313 có thể cơ bản tương tự như thành phần biến đổi tỷ lệ và lượng tử hóa 213. Các khối dư được biến đổi và được lượng tử hóa và các khối dự đoán tương ứng (theo dữ liệu điều khiển được

kết hợp) được chuyển tiếp đến thành phần mã hóa entropy 331 để mã hóa thành dòng bit. Thành phần mã hóa entropy 331 có thể cơ bản tương tự như thành phần CABAC và lấy khuôn dạng tiêu đề 231.

Các khối dư được biến đổi và được lượng tử hóa và/hoặc các khối dự đoán tương ứng cũng được chuyển tiếp từ thành phần biến đổi và lượng tử hóa 313 đến thành phần biến đổi ngược và lượng tử hóa 329 để khôi phục thành các khối tham chiếu cho việc sử dụng bởi thành phần bù chuyển động 321. Thành phần biến đổi ngược và lượng tử hóa 329 có thể cơ bản tương tự như thành phần biến đổi tỷ lệ và biến đổi ngược 229. Các bộ lọc trong vòng trong thành phần bộ lọc trong vòng 325 cũng được áp dụng cho các khối dư và/hoặc các khối tham chiếu được khôi phục, phụ thuộc vào ví dụ. Thành phần bộ lọc trong vòng 325 có thể cơ bản tương tự như thành phần phân tích điều khiển bộ lọc 227 và thành phần bộ lọc trong vòng 225. Thành phần bộ lọc trong vòng 325 có thể bao gồm nhiều bộ lọc như được đề cập đối với thành phần bộ lọc trong vòng 225. Các khối được lọc sau đó được lưu trữ trong thành phần bộ đệm ảnh được giải mã 323 để sử dụng như các khối tham chiếu bởi thành phần bù chuyển động 321. Thành phần bộ đệm ảnh được giải mã 323 có thể cơ bản tương tự như thành phần bộ đệm ảnh được giải mã 223.

FIG.4 là sơ đồ khối minh họa ví dụ bộ giải mã video 400. Bộ giải mã video 400 có thể được sử dụng để thực hiện các chức năng giải mã của hệ thống codec 200 và/hoặc thực hiện các bước 111, 113, 115, và/hoặc 117 của phương pháp thao tác 100. Bộ giải mã 400 thu dòng bit, ví dụ từ bộ mã hóa 300, và tạo ra tín hiệu video đầu ra được khôi phục dựa trên dòng bit để hiển thị cho người dùng.

Dòng bit được thu bởi thành phần giải mã entropy 433. Thành phần giải mã entropy 433 được tạo cấu hình để thực hiện cơ chế giải mã entropy, như là mã hóa CAVLC, CABAC, SBAC, PIPE, hoặc các kỹ thuật mã hóa entropy khác. Ví dụ, thành phần giải mã entropy 433 có thể sử dụng thông tin tiêu đề để cung cấp ngữ cảnh để diễn giải dữ liệu bổ sung được mã hóa như các từ mã trong dòng bit. Thông tin được giải mã bao gồm thông tin mong muốn bất kỳ để giải mã tín hiệu video, như là dữ liệu điều khiển chung, dữ liệu điều khiển bộ lọc, thông tin phân vùng, dữ liệu chuyển động, dữ liệu dự đoán, và các hệ số biến đổi được lượng tử hóa từ các khối dư. Các hệ số biến đổi được lượng tử hóa được chuyển tiếp đến thành phần biến đổi ngược và lượng tử hóa 429 cho việc khôi phục thành các khối dư. Thành phần biến đổi ngược và lượng tử hóa 429 có thể tương tự như thành phần biến đổi ngược và lượng tử hóa 329.

Các khối dư được khôi phục và/hoặc các khối dự đoán được chuyển tiếp đến thành phần dự đoán ảnh nội bộ 417 cho việc khôi phục thành các khối ảnh dựa trên các thao tác dự đoán nội bộ. Thành phần dự đoán ảnh nội bộ 417 có thể tương tự như thành phần ước lượng ảnh nội bộ 215 và thành phần dự đoán ảnh nội bộ 217. Cụ thể, thành phần dự đoán ảnh nội bộ 417 sử dụng các chế độ dự đoán để định vị khối tham chiếu trong khung và áp dụng khối dư cho kết quả để khôi phục các khối ảnh được dự đoán nội bộ. Các khối ảnh được khôi phục được dự đoán nội bộ và/hoặc các khối dư và dữ liệu dự đoán liên đới tương ứng được chuyển tiếp đến thành phần bộ đệm ảnh được giải mã 423 qua thành phần các bộ lọc trong vòng 425, mà có thể cơ bản tương tự như thành phần bộ đệm ảnh được giải mã 223 và thành phần các bộ lọc trong vòng 225, theo cách tương ứng. Thành phần các bộ lọc trong vòng 425 lọc các khối ảnh được khôi phục, các khối dư và/hoặc các khối dự đoán, và thông tin này được lưu trữ trong thành phần bộ đệm ảnh được giải mã 423. Các khối ảnh được khôi phục từ thành phần bộ đệm ảnh được giải mã 423 được chuyển tiếp đến thành phần bù chuyển động 421 dùng cho dự đoán liên đới. Thành phần bù chuyển động 421 có thể về cơ bản tương tự như thành phần ước lượng chuyển động 221 và/hoặc thành phần bù chuyển động 219. Cụ thể, 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 khôi phục khối ảnh. Việc dẫn đến các khối được khôi phục có thể cũng được chuyển tiếp qua thành phần bộ lọc trong vòng 425 đến thành phần bộ đệm ảnh được giải mã 423. Thành phần bộ đệm ảnh được giải mã 423 tiếp tục lưu trữ các khối ảnh được khôi phục bổ sung, mà có thể được khôi phục thành các khung qua thông tin phân vùng. Các khung này có thể cũng được đặt trong chuỗi. Chuỗi được xuất ra cho màn hình dưới dạng tín hiệu video đầu ra được khôi phục.

Theo phần nêu trên, các kỹ thuật nén video thực hiện việc dự đoán không gian (nội ảnh) và/hoặc thời gian (liên ảnh) để làm giảm hoặc loại bỏ dư thừa vốn có trong các chuỗi video. Đối với việc mã hóa video dựa trên khối, lát video (tức là ảnh video hoặc một phần của ảnh video) có thể được phân chia thành các khối video, mà cũng có thể được gọi là các khối cây, khối cây mã hóa (CTB), đơn vị cây mã hóa (CTU), đơn vị mã hóa (CU) và/hoặc các nút mã hóa. Các khối video trong lát được mã hóa nội bộ (I) của ảnh được mã hóa bằng cách sử dụng dự đoán không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng một ảnh. Các khối video trong lát được mã hóa liên đới (P hoặc B) của ảnh có thể sử dụng dự đoán không gian đối với các mẫu tham chiếu trong các khối lân cận trong cùng ảnh

hoặc dự đoán thời gian đối với các mẫu tham chiếu trong các ảnh tham chiếu khác. Các ảnh có thể được gọi là các khung, và các ảnh tham chiếu có thể được gọi là các khung tham chiếu.

Dự đoán theo không gian hoặc thời gian dẫn đến khối dự đoán đối với khối cần được mã hóa. Dữ liệu dư biểu diễn độ chênh lệch điểm ảnh giữa khối gốc cần được mã hóa và khối dự đoán. Khối được mã hóa liên đới được mã hóa theo vectơ chuyển động mà trỏ tới khối của các mẫu tham chiếu mà tạo thành khối dự đoán, và dữ liệu dư mà chỉ báo độ chênh lệch giữa khối được mã hóa và khối dự đoán. Khối được mã hóa nội bộ được mã hóa theo chế độ mã hóa nội bộ và dữ liệu dư. Để nén tiếp, dữ liệu dư có thể được biến đổi từ miền điểm ảnh sang miền biến đổi, mà dẫn đến các hệ số biến đổi dư, mà sau đó có thể được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa, được sắp xếp ban đầu trong mảng hai chiều, có thể được quét để tạo ra vectơ một chiều của các hệ số biến đổi và việc mã hóa entropy có thể được áp dụng để đạt được độ nén nhiều hơn.

Việc nén hình ảnh và video đã có sự phát triển nhanh chóng, mà dẫn đến các tiêu chuẩn mã hóa khác nhau. Các tiêu chuẩn mã hóa video này bao gồm ITU-T H.261, Tổ chức tiêu chuẩn hóa quốc tế/Ủy ban kỹ thuật điện quốc tế (ISO/IEC) MPEG-1 Phần 2, ITU-T H.262 hoặc ISO/IEC MPEG-2 Phần 2, ITU-T H.263, ISO/IEC MPEG-4 phần 2, mã hóa video cải tiến (AVC-Advanced Video Coding), còn được gọi là ITU-T H.264 hoặc ISO/IEC MPEG-4 Phần 10 và mã hóa video hiệu quả cao (HEVC-High Efficiency Video Coding), còn được gọi là ITU-T H.265 hoặc MPEG-H Phần 2. AVC bao gồm các phần mở rộng như Mã hóa video có thể biến đổi (SVC-Scalable Video Coding), Mã hóa video đa cảnh nhìn (MVC-Multiview Video Coding) và Mã hóa video đa cảnh nhìn cộng với độ sâu (MVC + D) và 3D AVC (3D-AVC). HEVC bao gồm các phần mở rộng như HEVC có thể biến đổi (SHVC), HEVC đa cảnh nhìn (MV-HEVC) và 3D HEVC (3D-HEVC).

Ngoài ra còn có tiêu chuẩn mã hóa video mới, được đặt tên là Mã hóa video đa năng (VVC-Versatile Video Coding), được phát triển bởi nhóm chuyên gia video liên kết (JVET) của ITU-T và ISO/IEC. Trong khi tiêu chuẩn VVC có một vài dự thảo hoạt động, một dự thảo hoạt động (WD-Working Draft) của VVC cụ thể là, B. Bross, J. Chen, và S. Liu, “Mã hóa video đa năng (Dự thảo 5),” JVET-N1001-v3, hội thảo JVET thứ 13, 27 tháng 3 năm 2019 (Dự thảo VVC 5) được tham chiếu ở đây.

Các mô hình phân chia ảnh trong HEVC được mô tả.

HEVC bao gồm bốn mô hình phân chia ảnh khác nhau, cụ thể là các lát thường, các lát phụ thuộc, các ô, và Xử lý song song mặt sóng (WPP-Wavefront Parallel Processing), mà có thể được áp dụng để so khớp kích cỡ đơn vị truyền lớn nhất (MTU-Maximum Transfer Unit), xử lý song song, và độ trễ đầu-cuối được làm giảm.

Các lát thường là tương tự như trong H.264/AVC. Mỗi lát thường được đóng gói trong chính đơn vị NAL và dự đoán nội ảnh (dự đoán mẫu nội bộ, dự đoán thông tin chuyển động, dự đoán chế độ mã hóa) và sự phụ thuộc mã hóa entropy qua các biên lát bị vô hiệu hóa. Do đó, lát thường có thể được khôi phục độc lập với các lát thường khác trong cùng ảnh (mặc dù vẫn có thể có sự phụ thuộc lẫn nhau do các hoạt động lọc vòng).

Lát thường là công cụ duy nhất mà có thể được sử dụng để song song hóa mà cũng khả dụng, dưới dạng gần như giống đồng nhất, trong H.264/AVC. Song song hóa dựa trên các lát thường không yêu cầu nhiều truyền thông liên bộ xử lý hoặc liên lõi (ngoại trừ việc chia sẻ dữ liệu liên bộ xử lý hoặc giữa liên lõi để bù chuyển động khi giải mã ảnh được mã hóa dự đoán, mà thường nặng hơn nhiều so với việc chia sẻ dữ liệu liên bộ xử lý hoặc liên lõi do việc dự đoán nội ảnh). Tuy nhiên, vì cùng lý do, việc sử dụng các lát thường có thể sinh ra thông tin tiêu đề mã hóa lớn do giá trị bit của thông tin tiêu đề lát và do thiếu sự dự đoán giữa các biên lát. Ngoài ra, các lát thường (ngược với các công cụ khác được đề cập dưới đây) cũng đóng vai trò là cơ chế chính để phân chia dòng bit để thỏa mãn các yêu cầu kích cỡ MTU, do sự độc lập nội ảnh của các lát thường và mỗi lát thường được đóng gói trong chính đơn vị lớp trừu tượng mạng (NAL-Network Adaptation Layer). Trong nhiều trường hợp, mục tiêu song song hóa và mục tiêu thỏa mãn kích thước MTU mâu thuẫn với yêu cầu đối với cấu trúc lát trong ảnh. Việc giải quyết tình trạng này đã dẫn đến sự phát triển của các công cụ song song hóa được đề cập dưới đây.

Các lát phụ thuộc có thông tin tiêu đề lát ngắn và cho phép phân chia dòng bit theo các biên khối cây mà không phá vỡ bất kỳ dự đoán ảnh nội bộ. Về cơ bản, các lát phụ thuộc cung cấp sự phân chia của các lát thường thành nhiều đơn vị NAL, để giảm độ trễ đầu-cuối bằng cách cho phép một phần của lát thường được gửi đi trước khi việc mã hóa toàn bộ lát thông thường kết thúc.

Trong WPP, ảnh được phân chia thành các hàng đơn lẻ của khối cây mã hóa (CTB-coding tree block). Dự đoán và giải mã entropy được phép sử dụng dữ

liệu từ CTB trong các phân vùng khác. Xử lý song song có thể được thực hiện thông qua giải mã song song các hàng CTB, trong đó việc bắt đầu giải mã hàng CTB bị làm trễ bởi hai CTB, do đó, để đảm bảo rằng dữ liệu liên quan đến CTB nêu trên và bên phải của CTB mục tiêu là khả dụng trước CTB mục tiêu đang được giải mã. Nhờ sử dụng khởi động so le này (mà xuất hiện như mặt sóng khi được biểu diễn bằng đồ thị), việc song song hóa có thể được thực hiện với tối đa số lượng bộ vi xử lý/lỗi như ảnh chứa các hàng CTB. Do việc dự đoán nội ảnh giữa các hàng khối cây lân cận trong ảnh được cho phép, việc truyền thông liên bộ xử lý/liên lỗi được yêu cầu để cho phép dự đoán ảnh nội bộ có thể là quan trọng. Việc phân chia WPP không dẫn đến việc tạo ra các đơn vị NAL bổ sung so với khi không được áp dụng, do đó WPP không phải là công cụ để so khớp kích cỡ MTU. Tuy nhiên, nếu việc so khớp kích cỡ MTU được yêu cầu, các lát thường có thể được sử dụng với WPP, với thông tin tiêu đề mã hóa nhất định.

Các ô xác định các biên ngang và dọc mà phân chia ảnh thành các cột và hàng ô. Thứ tự quét của các CTB được thay đổi thành cục bộ trong ô (theo thứ tự quét mảnh CTB của ô), trước khi giải mã CTB trên cùng bên trái của ô tiếp theo trong thứ tự quét mảnh ô của ảnh. Tương tự như các lát thường, các ô phá vỡ các phụ thuộc dự đoán nội ảnh cũng như các phụ thuộc giải mã entropy. Tuy nhiên, các ô không cần phải được gộp vào các đơn vị NAL riêng biệt (tương tự như WPP về vấn đề này); do đó không thể sử dụng các ô để so khớp kích cỡ MTU. Mỗi ô có thể được xử lý bởi một bộ xử lý/lỗi và việc truyền thông liên bộ xử lý/liên lỗi được yêu cầu để dự đoán ảnh nội bộ giữa các bộ xử lý mà giải mã các ô lân cận bị giới hạn ở việc truyền tải thông tin tiêu đề lát chia sẻ trong các trường hợp một lát kéo dài hơn một ô và việc lọc vòng liên quan đến chia sẻ các mẫu và siêu dữ liệu được khôi phục. Khi nhiều hơn một ô hoặc phân đoạn WPP được bao gồm trong lát, độ dịch byte điểm đầu vào đối với mỗi ô hoặc phân đoạn WPP khác với ô đầu tiên trong lát được báo hiệu trong thông tin tiêu đề lát.

Để đơn giản, các hạn chế đối với việc áp dụng bốn mô hình phân chia ảnh khác nhau đã được quy định trong HEVC. Chuỗi video được mã hóa định trước không thể bao gồm cả các ô và mặt sóng cho hầu hết các cấu hình được chỉ định trong HEVC. Đối với mỗi lát và ô, phải đáp ứng một trong hai hoặc cả hai điều kiện sau: 1) tất cả các khối cây được mã hóa trong lát thuộc cùng ô; 2) tất cả các khối cây được mã hóa trong ô đều thuộc về cùng lát. Cuối cùng, phân đoạn mặt

sóng chứa chính xác một hàng CTB và khi WPP được sử dụng, nếu lát bắt đầu trong hàng CTB, thì nó phải kết thúc trong cùng hàng CTB.

Các mô hình phân chia ảnh trong VVC được mô tả.

HEVC bao gồm bốn mô hình phân chia ảnh khác nhau, cụ thể là các lát, ô và viên, và xử lý song song mặt sóng (WPP), mà có thể được áp dụng để so khớp kích cỡ Đơn vị truyền lớn nhất (MTU), xử lý song song và giảm độ trễ đầu-cuối.

Các ô trong VVC cũng giống như các ô trong HEVC. Các ô xếp xác định ranh giới ngang và dọc để phân chia hình ảnh thành các cột và hàng ô. Trong VVC, khái niệm về các ô được cải tiến hơn nữa bằng cách cho phép ô được phân chia thêm theo chiều ngang để tạo thành các viên (brick). Ô không được phân chia thêm nữa cũng được coi là viên. Thứ tự quét của các CTB được thay đổi thành cục bộ trong viên (theo thứ tự quét mảnh CTB của viên), trước khi giải mã CTB trên cùng bên trái của viên tiếp theo trong thứ tự quét mảnh viên của ảnh.

Các lát trong VVC bao gồm một hoặc nhiều viên. Mỗi lát được đóng gói trong chính đơn vị NAL và dự đoán ảnh nội bộ (dự đoán mẫu nội bộ, dự đoán thông tin chuyển động, dự đoán chế độ mã hóa) và sự phụ thuộc mã hóa entropy qua các biên lát bị vô hiệu hóa. Do đó, lát thường có thể được khôi phục độc lập với các lát thường khác trong cùng ảnh (mặc dù vẫn có thể có sự phụ thuộc lẫn nhau do các hoạt động lọc vòng lặp). VVC định nghĩa hai loại lát là: lát hình chữ nhật và lát quét mảnh. Lát hình chữ nhật bao gồm một hoặc nhiều viên mà chiếm giữ vùng hình chữ nhật trong ảnh. Lát quét mảnh bao gồm một hoặc nhiều viên mà nằm trong thứ tự quét mảnh của các viên trong ảnh.

Đặc tính WPP trong VVC tương tự như đặc tính WPP trong HEVC với sự khác biệt là HEVC WPP có độ trễ hai (2) CTU trong khi VVC WPP có độ trễ một (1) CTU. Đối với HEVC WPP, dòng giải mã mới có thể bắt đầu giải mã CTU đầu tiên trong hàng CTU được gán của nó sau khi hàng CTU trước đó có hai CTU đầu tiên đã được giải mã; mặt khác, đối với VVC WPP, dòng giải mã mới có thể bắt đầu giải mã CTU đầu tiên trong hàng CTU được gán của nó sau khi hàng CTU trước đó đã giải mã CTU đầu tiên.

Báo hiệu của các ô, các viên và các lát trong PPS được mô tả.

Báo hiệu hiện tại của ô, viên và lát (đặc biệt là lát hình chữ nhật) trong PPS là như sau:

pic_parameter_set_rbsp() {	Nhãn mô tả
...	
single_tile_in_pic_flag	u(1)
if(!single_tile_in_pic_flag) {	
uniform_tile_spacing_flag	u(1)
if(uniform_tile_spacing_flag) {	
tile_cols_width_minus1	ue(v)
tile_rows_height_minus1	ue(v)
} else {	
num_tile_columns_minus1	ue(v)
num_tile_rows_minus1	ue(v)
for(i = 0; i < num_tile_columns_minus1; i++)	
tile_column_width_minus1[i]	ue(v)
for(i = 0; i < num_tile_rows_minus1; i++)	
tile_row_height_minus1[i]	ue(v)
}	
brick_splitting_present_flag	u(1)
if(uniform_tile_spacing_flag && brick_splitting_present_flag)	
num_tiles_in_pic_minus1	ue(v)
for(i = 0; brick_splitting_present_flag && i <= num_tiles_in_pic_minus1 + 1; i++) {	
if(RowHeight[i] > 1)	
brick_split_flag[i]	u(1)
if(brick_split_flag[i]) {	
if(RowHeight[i] > 2)	
uniform_brick_spacing_flag[i]	u(1)
if(uniform_brick_spacing_flag[i])	
brick_height_minus1[i]	ue(v)
else {	
num_brick_rows_minus2[i]	ue(v)
for(j = 0; j <= num_brick_rows_minus2[i]; j++)	
brick_row_height_minus1[i][j]	ue(v)
}	
}	
}	
single_brick_per_slice_flag	u(1)
if(!single_brick_per_slice_flag)	
rect_slice_flag	u(1)
if(rect_slice_flag && !single_brick_per_slice_flag) {	
num_slices_in_pic_minus1	ue(v)
bottom_right_brick_idx_length_minus1	ue(v)

for(i = 0; i < num_slices_in_pic_minus1; i++) {	
bottom_right_brick_idx_delta[i]	u(v)
brick_idx_delta_sign_flag[i]	u(1)
}	
}	
loop_filter_across_bricks_enabled_flag	u(1)
if(loop_filter_across_bricks_enabled_flag)	
loop_filter_across_slices_enabled_flag	u(1)
}	
if(rect_slice_flag) {	
signalled_slice_id_flag	u(1)
if(signalled_slice_id_flag) {	
signalled_slice_id_length_minus1	ue(v)
for(i = 0; i <= num_slices_in_pic_minus1; i++)	
slice_id[i]	u(v)
}	
}	
...	
}	

Các lát được kết hợp với thông tin tiêu đề của lát mà chứa phần tử cú pháp để mô tả thuộc tính của các lát mà cần thiết để giải mã các lát. Ví dụ về bảng cú pháp thông tin tiêu đề lát và phần giới thiệu của ngữ nghĩa thông tin tiêu đề lát được đưa ra dưới đây.

Cú pháp thông tin tiêu đề lát chung.

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

if(nal_unit_type == IDR_W_RADL nal_unit_type == IDR_N_LP nal_unit_type == CRA_NUT NalUnitType == GDR_NUT)	
no_output_of_prior_pics_flag	u(1)
if(output_flag_present_flag)	
pic_output_flag	u(1)
...	
}	

Ngữ nghĩa thông tin tiêu đề lát cắt chung.

Khi được hiện diện, giá trị của mỗi phần tử cú pháp thông tin tiêu đề lát slice_pic_parameter_set_id, non_reference_picture_flag, colour_plane_id, slice_pic_order_cnt_lsb, recovery_poc_cnt, no_output_of_prior_pics_flag, pic_output_flag, và slice_temporal_mvp_enabled_flag sẽ giống nhau trong tất cả các thông tin tiêu đề lát của ảnh được mã hóa.

Biến CuQpDeltaVal, chỉ rõ độ chênh lệch giữa tham số lượng tử hóa độ chói đối với đơn vị mã hóa mà chứa cu_qp_delta_abs và dự đoán của nó, được thiết lập bằng 0. Các biến CuQpOffsetCb, CuQpOffsetCr và CuQpOffsetCbCr, mà chỉ rõ các giá trị được sử dụng khi xác định các giá trị tương ứng của các tham số lượng tử hóa Qp'Cb, Qp'Cr và Qp'CbCr đối với đơn vị mã hóa mà chứa cu_chroma_qp_offset_flag, đều được thiết lập bằng 0.

Dấu phân cách đơn vị truy nhập được mô tả.

Tại hội thảo JVET lần thứ 15 ở Gothenburg, đã được đồng ý phải bắt buộc có sự hiện diện của dấu phân cách đơn vị truy nhập (AUD) cho mỗi ảnh trong các dòng bit VVC. AUD bắt buộc phải có mặt trước lát thứ nhất của mỗi ảnh, mà giúp bộ giải mã phát hiện sự bắt đầu của ảnh mới trong dòng bit VVC.

Cú pháp và ngữ nghĩa của AUD RBSP là như sau.

access_unit_delimiter_rbsp() {	Nhãn mô tả
pic_type	u(3)
rbsp_trailing_bits()	
}	

Dấu phân cách đơn vị truy nhập được sử dụng để chỉ báo sự bắt đầu của đơn vị truy nhập và loại lát có trong các ảnh được mã hóa trong đơn vị truy nhập

mà chứa đơn vị NAL dấu phân cách đơn vị truy nhập. Không có xử lý giải mã quy chuẩn được kết hợp với dấu phân cách đơn vị truy nhập.

pic_type chỉ báo rằng các giá trị slice_type đối với tất cả các lát của ảnh được mã hóa trong đơn vị truy nhập mà chứa đơn vị NAL dấu phân cách đơn vị truy nhập là các thành viên của tập hợp được liệt kê trong Bảng 7-3 cho giá trị định trước của pic_type. Giá trị của pic_type sẽ bằng 0, 1, hoặc 2 trong các dòng bit mà tuân theo phiên bản của tiêu chuẩn kỹ thuật này. Các giá trị khác của pic_type được dành riêng cho sử dụng trong tương lai bởi ITU T | ISO/IEC. Các bộ giải mã mà tuân theo phiên bản này của tiêu chuẩn kỹ thuật này sẽ bỏ qua các giá trị dành riêng của pic_type.

Bảng 7-3 - Diễn giải của tham số pic_type

pic_type	các giá trị slice_type có thể xuất hiện trong ảnh được mã hóa
0	I
1	P, I
2	B, P, I

Các vấn đề với dấu phân cách đơn vị truy nhập được mô tả.

Bằng cách bắt buộc phải có AUD cho mỗi ảnh, một số vấn đề được xác định như sau.

Chỉ có một dấu phân cách đơn vị truy nhập được hiện diện cho mỗi ảnh hoặc đơn vị truy nhập và được giả định nằm trong đơn vị NAL mà là đầu tiên trong đơn vị truy nhập. Tuy nhiên, nếu đơn vị truy nhập chứa nhiều hơn một ảnh, chẳng hạn như trong các trường hợp dòng bit đa lớp, thì dấu phân cách đơn vị truy nhập chỉ có thể hỗ trợ phát hiện mỗi đơn vị truy nhập mà không phải mỗi ảnh.

Trong dòng bit đơn lớp, do mỗi ảnh có một AUD, một số phần tử cú pháp mà đang được báo hiệu trong thông tin tiêu đề lát nhưng bị ràng buộc là giống nhau đối với tất cả các lát của cùng ảnh sẽ được chuyển sang AUD. Mỗi phần tử cú pháp có thể được gọi là các phần tử cú pháp mức ảnh. Tuy nhiên, do đơn vị NAL mà chứa AUD được cho là đơn vị NAL đầu tiên trong đơn vị truy nhập, đơn vị NAL mà chứa AUD có thể đứng trước các đơn vị NAL mà chứa các tập hợp tham số như SPS, PPS, v.v. Thứ tự của AUD và các tập hợp tham số giới hạn tính linh hoạt trong việc di chuyển các phần tử cú pháp mức ảnh từ thông tin tiêu đề lát sang AUD và việc phân tích các phần tử cú pháp này đó có thể phụ thuộc vào thông tin được báo hiệu trong các tập hợp tham số.

Mô tả về các kỹ thuật được bộc lộ ở đây dựa trên tiêu chuẩn mã hóa video đang được phát triển là Mã hóa video đa năng (VVC) của nhóm chuyên gia video liên kết (JVET) của ITU-T và ISO/IEC. Tuy nhiên, các kỹ thuật này cũng áp dụng tới các tiêu chuẩn kỹ thuật mã hóa-giải mã video khác.

Như đã lưu ý ở trên, AUD đã được sử dụng để chỉ báo sự bắt đầu của mỗi AU, và do đó là điểm bắt đầu của ảnh, khi dòng bit chứa một lớp. Tuy nhiên, AUD không thể cung cấp chức năng tương tự cho dòng bit đa lớp vì dòng bit đa lớp bao gồm nhiều ảnh tại các độ phân giải khác nhau trong cùng AU. Do đó, khi đơn vị truy nhập chứa nhiều hơn một ảnh, chẳng hạn như trong dòng bit nhiều lớp, AUD chỉ có thể chỉ báo điểm bắt đầu của AU, nhưng không thể chỉ báo điểm bắt đầu của cụ thể trong AU.

Phần bộc lộ ở đây là các kỹ thuật mà đảm bảo rằng đơn vị ảnh (PU) chứa chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh, mà cho phép ký hiệu nhận dạng (ID) lớp của đơn vị lớp trừu tượng mạng (NAL) bằng với ID lớp của đơn vị NAL lớp mã hóa video (VCL) được liên kết và ID theo thời gian của đơn vị NAL bằng với ID theo thời gian của đơn vị ảnh (PU) bao gồm đơn vị NAL khi đơn vị NAL có kiểu đơn vị NAL thông tin tiêu đề ảnh (PH_NUT). Việc sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, bộ mã hóa/bộ giải mã (tức là, “bộ mã hóa-giải mã (codec)”) trong mã hóa video được cải thiện so với các bộ mã hóa-giải mã hiện tại. Theo vấn đề thực tế, xử lý mã hóa video được cải thiện mang lại cho người dùng sự trải nghiệm người dùng tốt hơn khi các video được gửi, được thu, và/hoặc được quan sát.

Khả năng biến đổi trong mã hóa video thường được hỗ trợ bằng cách sử dụng các kỹ thuật mã hóa đa lớp. Dòng bit đa lớp bao gồm một lớp cơ sở (BL-base layer) và một hoặc nhiều lớp nâng cao (EL-enhancement layer). Ví dụ về khả năng biến đổi bao gồm khả năng biến đổi không gian, khả năng biến đổi chất lượng/tín hiệu trên tạp âm (SNR-signal-to-noise), khả năng biến đổi đa cảnh nhìn, v.v. Khi sử dụng kỹ thuật mã hóa đa lớp, ảnh hoặc một phần của ảnh có thể được mã hóa (1) mà không sử dụng ảnh tham chiếu, tức là sử dụng dự đoán nội bộ; (2) bằng cách tham chiếu đến các ảnh tham chiếu nằm trong cùng lớp, tức là sử dụng dự đoán liên đới; hoặc (3) bằng cách tham chiếu đến các ảnh tham chiếu nằm trong (các) lớp khác, tức là sử dụng dự đoán lớp liên đới. Ảnh tham chiếu được sử dụng để dự đoán

liên lớp của ảnh hiện tại được gọi là ảnh tham chiếu liên lớp (ILRP-inter-layer reference picture).

FIG.5 là sơ đồ giản lược minh họa ví dụ về dự đoán dựa trên lớp 500, ví dụ như được thực hiện để xác định các MV tại bước nén khối 105, bước giải mã khối 113, bộ ước lượng chuyển động 221, bộ bù chuyển động 219, bộ bù chuyển động 321, và/hoặc bộ bù chuyển động 421. Dự đoán dựa trên lớp 500 tương thích với dự đoán liên đới một chiều và/hoặc dự đoán liên đới hai chiều, nhưng cũng được thực hiện giữa các ảnh trong các lớp khác nhau.

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

Các ảnh 511-518 trong các lớp khác nhau 531-532 được tạo cấu hình để được hiển thị thay thế. Do đó, các ảnh 511-518 trong các lớp khác nhau 531-532 có thể chia sẻ cùng ký hiệu nhận dạng (ID) thời gian khi các ảnh được bao gồm trong cùng AU. Như được sử dụng ở đây, AU là tập hợp của một hoặc ảnh được mã hóa có cùng thời gian hiển thị để xuất ra từ DPB. Ví dụ, bộ giải mã có thể giải mã và hiển thị hình ảnh 515 tại thời điểm hiển thị hiện tại nếu muốn có ảnh nhỏ hơn hoặc bộ giải mã có thể giải mã và hiển thị hình ảnh 511 tại thời điểm hiển thị hiện tại nếu muốn có ảnh lớn hơn. Do đó, các hình ảnh 511-514 tại lớp cao hơn $N+1$ 532 về bản chất chứa cùng dữ liệu ảnh với các ảnh tương ứng 515-518 tại lớp

thấp hơn N 531 (bất chấp độ chênh lệch về kích cỡ ảnh). Cụ thể, ảnh 511 về bản chất chứa dữ liệu ảnh giống như ảnh 515, ảnh 512 chứa dữ liệu hình ảnh về bản chất giống như ảnh 516, v.v.

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

Các ảnh 511-518 cũng có thể được mã hóa bằng cách tham chiếu đến các ảnh khác 511-518 trong các lớp khác nhau. Xử lý này được gọi là dự đoán lớp liên đới 521 và được thể hiện bằng các mũi tên đứt nét. Dự đoán lớp liên đới 521 là cơ chế mã hóa các mẫu của ảnh hiện tại bằng cách tham chiếu đến các mẫu được chỉ báo trong ảnh tham chiếu trong đó ảnh hiện tại và ảnh tham chiếu nằm ở các lớp khác nhau và do đó có các ID lớp khác nhau. Ví dụ, hình ảnh ở lớp thấp hơn N 531 có thể được sử dụng làm ảnh tham chiếu để mã hóa ảnh tương ứng tại lớp cao hơn N+1 532. Theo ví dụ cụ thể, ảnh 511 có thể được mã hóa bằng cách tham chiếu đến ảnh 515 theo dự đoán lớp liên đới 521. Trong trường hợp này, ảnh 515 được sử dụng làm ảnh tham chiếu lớp liên đới. Ảnh tham chiếu lớp liên đới là ảnh tham chiếu được sử dụng cho dự đoán lớp liên đới 521. Trong hầu hết các trường hợp, dự đoán lớp liên đới 521 bị hạn chế sao cho ảnh hiện tại, chẳng hạn như ảnh 511,

chỉ có thể sử dụng các ảnh tham chiếu lớp liên đới được bao gồm trong cùng AU và tại lớp thấp hơn, chẳng hạn như ảnh 515. Khi nhiều lớp (ví dụ, nhiều hơn hai) là khả dụng, dự đoán lớp liên đới 521 có thể mã hóa/giải mã ảnh hiện tại dựa trên (các) ảnh tham chiếu lớp liên đới ở các mức thấp hơn so với ảnh hiện tại.

Bộ mã hóa video có thể sử dụng dự đoán dựa trên lớp 500 để mã hóa các ảnh 511-518 thông qua nhiều kết hợp và/hoặc hoán vị khác nhau của dự đoán liên đới 523 và dự đoán lớp liên đới 521. Ví dụ, ảnh 515 có thể được mã hóa theo dự đoán nội bộ. Các ảnh 516-518 sau đó có thể được mã hóa theo dự đoán liên đới 523 bằng cách sử dụng ảnh 515 như là ảnh tham chiếu. Hơn nữa, ảnh 511 có thể được mã hóa theo dự đoán lớp liên đới 521 bằng cách sử dụng ảnh 515 như là ảnh tham chiếu lớp liên đới. Các ảnh 512-514 sau đó có thể được mã hóa theo dự đoán liên đới 523 bằng cách sử dụng ảnh 511 như là ảnh tham chiếu. Do đó, ảnh tham chiếu có thể đóng vai trò là cả ảnh tham chiếu lớp đơn và ảnh tham chiếu lớp liên đới cho các cơ chế mã hóa khác nhau. Bằng cách mã hóa ảnh lớp cao hơn N+1 532 dựa trên ảnh lớp thấp hơn N 531, lớp cao hơn N+1 532 có thể tránh sử dụng dự đoán nội bộ, mà có hiệu quả mã hóa thấp hơn nhiều so với dự đoán liên đới 523 và dự đoán lớp liên đới 521. Do đó, hiệu quả mã hóa kém của dự đoán nội bộ có thể bị giới hạn ở những ảnh chất lượng thấp nhất/nhỏ nhất và do đó bị giới hạn trong việc mã hóa lượng dữ liệu video nhỏ nhất. Các ảnh được sử dụng làm ảnh tham chiếu và/hoặc các ảnh tham chiếu lớp liên đới có thể được chỉ báo trong các mục của (các) danh sách ảnh tham chiếu được chứa trong cấu trúc danh sách ảnh tham chiếu.

Mỗi AU 506 trong FIG.5 có thể chứa một số ảnh. Ví dụ, một AU 506 có thể chứa các ảnh 511 và 515. AU 506 khác có thể chứa các ảnh 512 và 516. Do vậy, mỗi AU 506 là tập hợp của một hoặc nhiều ảnh được mã hóa được kết hợp với cùng thời gian hiển thị (ví dụ, cùng ID theo thời gian) để xuất ra từ bộ đệm ảnh được giải mã (DPB) (ví dụ, để hiển thị tới người dùng) Mỗi AUD 508 là chỉ báo hoặc cấu trúc dữ liệu được sử dụng để chỉ báo sự bắt đầu của AU (ví dụ, AU 506) hoặc biên giữa các AU.

Các họ mã hóa video H.26x trước đây đã hỗ trợ khả năng biến đổi trong (các) cấu hình riêng biệt với (các) cấu hình để mã hóa đơn lớp. Mã hóa video có thể biến đổi (SVC) là phần mở rộng có thể biến đổi của AVC/H.264 mà cung cấp hỗ trợ cho khả năng biến đổi theo không gian, thời gian và chất lượng. Đối với SVC, cờ được báo hiệu trong mỗi khối macro (MB-macroblock) trong các ảnh EL để chỉ báo rằng EL MB có được dự đoán bằng cách sử dụng khối được sắp xếp

cùng vị trí từ lớp thấp hơn hay không. Việc dự đoán từ khối được sắp xếp cùng vị trí có thể bao gồm kết cấu, vector chuyển động và/hoặc các chế độ mã hóa. Các cách thức thực hiện của SVC không thể sử dụng lại trực tiếp các cách thức thực hiện H.264/AVC không cải biến trong thiết kế của chúng. Cú pháp và xử lý giải mã các khối macro SVC EL khác với cú pháp và xử lý giải mã H.264/AVC.

HEVC có thể biến đổi (SHVC) là phần mở rộng của tiêu chuẩn HEVC/H.265 mà cung cấp hỗ trợ cho khả năng biến đổi theo không gian và chất lượng, HEVC đa cảnh nhìn (MV-HEVC) là phần mở rộng của HEVC/H.265 mà cung cấp hỗ trợ cho khả năng biến đổi đa cảnh nhìn và 3D HEVC (3D-HEVC) là phần mở rộng của HEVC/H.264 mà cung cấp hỗ trợ mã hóa video ba chiều (3D) tiên tiến hơn và hiệu quả hơn so với MV-HEVC. Lưu ý rằng khả năng biến đổi theo thời gian được bao gồm như là phần toàn vẹn của mã hóa-giải mã HEVC đơn lớp. Thiết kế của phần mở rộng đa lớp của HEVC sử dụng ý tưởng trong đó các ảnh được giải mã được sử dụng để dự đoán lớp liên đới chỉ đến từ cùng đơn vị truy nhập (AU) và được xem là ảnh tham chiếu dài hạn (LTRP) và được gán các chỉ số tham chiếu trong (các) danh sách ảnh tham chiếu cùng với các ảnh tham chiếu theo thời gian khác trong lớp hiện tại. Dự đoán lớp liên đới (ILP) đạt được ở mức đơn vị dự đoán bằng cách thiết lập giá trị của chỉ số tham chiếu để tham chiếu đến (các) ảnh tham chiếu lớp liên đới trong (các) danh sách ảnh tham chiếu.

Đáng chú ý, cả hai đặc tính lấy mẫu lại ảnh tham chiếu và khả năng biến đổi không gian đều yêu cầu lấy mẫu lại ảnh tham chiếu hoặc một phần của nó. Lấy mẫu lại ảnh tham chiếu (RPR-Reference picture resampling) có thể được thực hiện tại mức ảnh hoặc mức khối mã hóa. Tuy nhiên, khi RPR được gọi là đặc tính mã hóa, đây là đặc tính để mã hóa lớp đơn. Mặc dù vậy, có thể hoặc thậm chí tốt hơn từ điểm thiết kế mã hóa-giải mã để sử dụng cùng một bộ lọc lấy mẫu lại cho cả đặc tính RPR của mã hóa lớp đơn và đặc tính khả năng biến đổi không gian cho mã đa lớp.

FIG.6 minh họa dòng bit video 650 có cấu trúc để thực hiện kỹ thuật làm mới giải mã từng bậc (GDR-gradual decoding refresh) 600. Như được sử dụng ở đây, dòng bit video 650 cũng có thể được gọi là dòng bit video được mã hóa, dòng bit hoặc các biến thể của chúng. Như được thể hiện trong FIG.6, dòng bit 650 bao gồm tập hợp tham số chuỗi (SPS) 652, tập hợp tham số ảnh (PPS) 654, thông tin tiêu đề lát 656 và dữ liệu ảnh 658.

SPS 652 chứa dữ liệu chung đối với tất cả ảnh trong chuỗi của các ảnh (SOP). Ngược lại, PPS 654 chứa dữ liệu chung đối với toàn bộ ảnh. Thông tin tiêu đề lát 656 chứa thông tin về lát hiện tại, chẳng hạn như loại lát, ảnh tham chiếu nào sẽ được sử dụng, v.v. SPS 652 và PPS 654 có thể được gọi chung là tập hợp tham số. SPS 652, PPS 654 và thông tin tiêu đề lát 656 là các loại đơn vị Lớp trừu tượng mạng (NAL-Network Abstraction Layer). Đơn vị NAL là cấu trúc cú pháp mà chứa chỉ báo về loại dữ liệu cần tuân theo (ví dụ, dữ liệu video được mã hóa) Các đơn vị NAL được phân loại thành lớp mã hóa video (VCL) và các đơn vị NAL không phải VCL. Các đơn vị VCL NAL chứa dữ liệu mà biểu diễn các giá trị của các mẫu trong các ảnh video và các đơn vị NAL không phải VCL chứa bất kỳ thông tin bổ sung được kết hợp như các tập hợp tham số (dữ liệu quan trọng mà có thể áp dụng tới một số đơn vị VCL NAL) và thông tin nâng cao bổ sung (thông tin thời gian và dữ liệu bổ sung khác có thể nâng cao khả năng sử dụng của tín hiệu video được giải mã nhưng không cần thiết để giải mã các giá trị của các mẫu trong các ảnh video). Người có hiểu biết trung bình trong lĩnh vực hiểu rằng dòng bit 650 có thể chứa các tham số và thông tin khác trong các ứng dụng thực tế.

Dữ liệu ảnh 658 của FIG.6 bao gồm dữ liệu được kết hợp với các ảnh hoặc video được mã hóa hoặc giải mã. Dữ liệu ảnh 658 có thể được gọi đơn giản là tải trọng hoặc dữ liệu đang được mang trong dòng bit 650. Theo phương án của sáng chế, dữ liệu ảnh 658 bao gồm CVS 608 (hoặc CLVS) mà chứa hình ảnh GDR 602, một hoặc nhiều ảnh kế tiếp 604, và ảnh điểm khôi phục 606. Theo phương án của sáng chế, ảnh GDR 602 được gọi là ảnh bắt đầu CVS (CVSS). CVS 608 là chuỗi video được mã hóa cho mọi chuỗi video lớp được mã hóa (CLVS) trong dòng bit video 650. Đáng chú ý, CVS và CLVS là giống nhau khi dòng bit video 650 bao gồm một lớp. CVS và CLVS chỉ khác nhau khi dòng bit video 650 bao gồm nhiều lớp. Theo phương án của sáng chế, các ảnh kế tiếp 604 có thể được coi là dạng của ảnh GDR vì các ảnh kế tiếp đứng trước ảnh điểm khôi phục 606 trong chu kỳ GDR.

Theo phương án của sáng chế, ảnh GDR 602, các ảnh kế tiếp 604 và ảnh điểm khôi phục 606 có thể xác định chu kỳ GDR trong CVS 608. Trong phương án, thứ tự giải mã bắt đầu với ảnh GDR 602, tiếp tục với các ảnh theo sau 604, và sau đó tiến hành đến ảnh khôi phục 606.

CVS 608 là một chuỗi các ảnh (hoặc các phần của nó) bắt đầu với ảnh GDR 602 và bao gồm tất cả các ảnh (hoặc các phần của nó) cho tới, nhưng không bao gồm, ảnh GDR tiếp theo hoặc cho đến cuối của dòng bit. Chu kỳ GDR là một

loạt các ảnh bắt đầu với ảnh GDR 602 và bao gồm tất cả các hình ảnh cho đến và bao gồm ảnh điểm khôi phục 606. Xử lý giải mã đối với CVS 608 luôn bắt đầu tại ảnh GDR 602.

Như được thể hiện trong FIG.6, kỹ thuật GDR 600 hoặc nguyên tắc hoạt động trên một loạt các ảnh bắt đầu với ảnh GDR 602 và kết thúc bằng ảnh điểm khôi phục 606. Ảnh GDR 602 chứa một vùng làm mới/sạch 610 mà chứa các khối có tất cả được mã hóa bằng cách sử dụng dự đoán nội bộ (tức là các khối được dự đoán nội bộ) và vùng chưa được làm mới/xấu 612 mà chứa các khối đã được mã hóa bằng cách sử dụng dự đoán liên đới (tức là các khối được dự đoán liên đới).

Ảnh kế tiếp 604 ngay cạnh ảnh GDR 602 chứa vùng 610 được làm mới/sạch có phần thứ nhất 610A được mã hóa bằng cách sử dụng dự đoán nội bộ và phần thứ hai 610B được mã hóa bằng cách sử dụng dự đoán liên đới. Phần thứ hai 610B được mã hóa bằng cách tham chiếu đến vùng được làm mới/sạch 610 của, ví dụ, ảnh trước đó trong chu kỳ GDR của CVS 608. Như được thể hiện, vùng được làm mới/sạch 610 của các ảnh kế tiếp 604 mở rộng khi xử lý mã hóa di chuyển hoặc tiến triển theo hướng nhất quán (ví dụ, từ trái sang phải), tương ứng thu nhỏ vùng chưa được làm mới/cũ 612. Cuối cùng, ảnh điểm khôi phục 606, chỉ chứa vùng được làm mới/sạch 610, thu được từ xử lý mã hóa. Đáng chú ý, và như sẽ được mô tả thêm bên dưới, phần thứ hai 610B của vùng được làm mới/sạch 610, được mã hóa như là các khối được dự đoán liên đới, chỉ có thể tham chiếu đến vùng được làm mới/sạch 610 trong ảnh tham chiếu.

Như được thể hiện trong FIG.6, ảnh GDR 602, ảnh kế tiếp 604 và ảnh điểm khôi phục 606 trong CVS 608 đều được chứa trong chính đơn vị VCL NAL 630 của chúng. Đơn vị VCL NAL 630 trong CVS 608 có thể được gọi là đơn vị truy nhập.

Theo phương án của sáng chế, đơn vị VCL NAL 630 mà chứa ảnh GDR 602 trong CVS 608 có loại đơn vị GDR NAL (GDR_NUT). Tức là, theo phương án của sáng chế, đơn vị VCL NAL 630 mà chứa ảnh GDR 602 trong CVS 608 có loại đơn vị NAL duy nhất của riêng nó so với ảnh kế tiếp 604 và ảnh điểm khôi phục 606. Trong phương án của sáng chế, GDR_NUT cho phép dòng bit 650 bắt đầu với hình ảnh GDR 602 thay vì dòng bit 650 phải bắt đầu bằng ảnh điểm truy cập ngẫu nhiên nội bộ (IRAP-intra random access point). Việc chỉ định đơn vị VCL NAL 630 của ảnh GDR 602 là GDR_NUT có thể chỉ báo, ví dụ, bộ giải mã rằng đơn vị VCL NAL khởi tạo 630 trong CVS 608 chứa ảnh GDR 602. Theo phương

án của sáng chế, ảnh GDR 602 là ảnh khởi tạo trong CVS 608. Theo phương án của sáng chế, ảnh GDR 602 là ảnh khởi tạo trong chu kỳ GDR.

FIG.7 là sơ đồ minh họa tìm kiếm chuyển động không mong muốn 700 khi sử dụng hạn chế của bộ mã hóa để hỗ trợ GDR. Như được thể hiện, tìm kiếm chuyển động 700 thể hiện ảnh hiện tại 702 và ảnh tham chiếu 704. Mỗi ảnh hiện tại 702 và ảnh tham chiếu 704 bao gồm vùng được làm mới 706 được mã hóa với dự đoán nội bộ, vùng được làm mới và được mã hóa với dự đoán liên đới và vùng chưa được làm mới 710. Vùng được làm mới 706, vùng được làm mới 708 và vùng chưa được làm mới 710 là tương tự với phần thứ nhất 610A của vùng được làm mới/sạch 610, phần thứ hai 610B của vùng được làm mới/sạch 610 và vùng không vùng làm mới/cũ 612 trong FIG.6.

Trong xử lý tìm kiếm chuyển động 700, bộ mã hóa bị hạn chế hoặc ngăn cản việc lựa chọn bất kỳ vectơ chuyển động nào 712 mà dẫn đến một số mẫu của khối tham chiếu 714 nằm ngoài vùng được làm mới 706. Điều này xảy ra ngay cả khi khối tham chiếu 714 cung cấp tiêu chuẩn giá trị méo tỷ lệ tốt nhất khi dự đoán khối hiện tại 716 trong ảnh hiện tại 702. Do đó, FIG.7 minh họa lý do không tối ưu trong tìm kiếm chuyển động 700 khi sử dụng hạn chế của bộ mã hóa để hỗ trợ GDR.

FIG.8 minh họa phương án về dòng bit video 800. Như được sử dụng ở đây, dòng bit video 800 cũng có thể được gọi là dòng bit video được mã hóa, dòng bit hoặc các biến thể của chúng. Như được thể hiện trong FIG.8, dòng bit 800 bao gồm ít nhất một đơn vị ảnh (PU) 801. Trong khi ba trong số các PU 801 được thể hiện trong FIG.8, số lượng của PU 801 khác nhau có thể có trong dòng bit 800 trong các ứng dụng thực tế. Mỗi PU 801 là tập hợp của các đơn vị NAL mà được kết hợp với nhau theo quy tắc phân loại cụ thể, là liên tiếp theo thứ tự giải mã và chứa chính xác một ảnh được mã hóa (ví dụ, ảnh 814). Trong một phương án, mỗi PU 801 được liên kết với ID theo thời gian 819.

Theo phương án của sáng chế, mỗi PU 801 chứa một hoặc nhiều thông tin sau: thông tin khả năng giải mã (DCI-decoding capability information) 802, tập hợp tham số video (VPS-video parameter set) 804, tập hợp tham số chuỗi (SPS-sequence parameter set) 806, tập hợp tham số ảnh (PPS-picture parameter set) 808, thông tin tiêu đề ảnh (PH-picture header) 812 và ảnh 814. Mỗi DCI 802, VPS 804, SPS 806 và PPS 808 có thể được gọi chung là tập hợp tham số. Trong phương án của sáng chế, các tập hợp tham số khác không được thể hiện trên FIG.8 có thể cũng

được chứa trong dòng bit 800 như, ví dụ, tập hợp tham số thích ứng (APS-adaption parameter set), mà là cấu trúc cú pháp mà chứa các phần tử cú pháp mà áp dụng tới 0 lát hoặc nhiều hơn như được xác định bởi 0 phần tử cú pháp hoặc nhiều hơn được tìm thấy trong các thông tin tiêu đề lát.

DCI 802, mà có thể cũng được gọi là tập hợp tham số giải mã (DPS-decoding parameter set) hoặc tập hợp tham số bộ giải mã, là cấu trúc cú pháp mà chứa các phần tử cú pháp mà áp dụng tới toàn bộ dòng bit. DCI 802 bao gồm các tham số mà cố định trong thời gian tồn tại của dòng bit video (ví dụ, dòng bit 800), mà có thể dịch thành thời gian tồn tại của phiên. DCI 802 có thể bao gồm cấu hình, mức, và thông tin cấu hình con để xác định điểm tương tác có độ phức tạp tối đa mà được đảm bảo không bao giờ vượt quá, ngay cả nếu việc nối các chuỗi video xảy ra trong phiên. Một cách tùy chọn DCI 802 còn bao gồm các cờ điều kiện, mà chỉ báo rằng dòng bit video sẽ bị hạn chế việc sử dụng các đặc tính nhất định như được chỉ báo bởi các giá trị của các cờ này. Với điều này, dòng bit có thể được gắn nhãn là không sử dụng các công cụ nhất định, mà cho phép cấp phát tài nguyên trong số những thứ khác trong cách thức thực hiện bộ giải mã. Tương tự tất cả các tập hợp tham số, DCI 802 được hiện diện khi được tham chiếu lần đầu tiên và được tham chiếu bởi ảnh đầu tiên trong chuỗi video, ngụ ý rằng nó phải được gửi giữa các đơn vị NAL thứ nhất trong dòng bit. Trong khi nhiều DCI 802 có thể nằm trong dòng bit, giá trị của các phần tử cú pháp trong đó không thể không nhất quán khi được tham chiếu.

VPS 804 bao gồm phần phụ thuộc giải mã hoặc thông tin để xây dựng tập hợp ảnh tham chiếu của các lớp nâng cao. VPS 804 cung cấp khía cạnh tổng thể hoặc của chuỗi có thể biến đổi, mà bao gồm những loại điểm hoạt động nào được cung cấp, cấu hình, bậc và mức của các điểm hoạt động và một số thuộc tính mức cao khác của dòng bit mà có thể được sử dụng như là cơ sở để đàm phán phiên và lựa chọn nội dung, v.v.

SPS 806 chứa dữ liệu chung đối với tất cả ảnh trong chuỗi của các ảnh (SOP). SPS 806 là cấu trúc cú pháp mà chứa các phần tử cú pháp mà áp dụng tới 0 hoặc nhiều CLVS như được xác định bởi nội dung của phần tử cú pháp được tìm thấy trong PPS được tham chiếu bởi phần tử cú pháp được tìm thấy trong mỗi thông tin tiêu đề ảnh. Ngược lại, PPS 808 chứa dữ liệu chung đối với toàn bộ ảnh. PPS 808 là cấu trúc cú pháp mà chứa các phần tử cú pháp mà áp dụng tới 0 hoặc nhiều

ảnh được mã hóa như được xác định bởi phần tử cú pháp được tìm thấy trong mỗi thông tin tiêu đề ảnh (ví dụ, PH 812).

DCI 802, VPS 804, SPS 806 và PPS 808 được chứa trong các loại đơn vị Lớp trừu tượng mạng (NAL) khác nhau. Đơn vị NAL là cấu trúc cú pháp mà chứa chỉ báo về loại dữ liệu cần tuân theo (ví dụ, dữ liệu video được mã hóa) Các đơn vị NAL được phân loại thành lớp mã hóa video (VCL) và các đơn vị NAL không phải VCL. Các đơn vị VCL NAL chứa dữ liệu mà biểu diễn các giá trị của các mẫu trong các ảnh video và các đơn vị NAL không phải VCL chứa bất kỳ thông tin bổ sung được kết hợp như các tập hợp tham số (dữ liệu quan trọng mà có thể áp dụng tới một số đơn vị VCL NAL) và thông tin nâng cao bổ sung (thông tin thời gian và dữ liệu bổ sung khác có thể nâng cao khả năng sử dụng của tín hiệu video được giải mã nhưng không cần thiết để giải mã các giá trị của các mẫu trong các ảnh video).

Theo phương án của sáng chế, DCI 802 được chứa trong đơn vị NAL không phải VCL được chỉ định là đơn vị DCI NAL hoặc đơn vị DPS NAL. Tức là, đơn vị DCI NAL có loại đơn vị DCI NAL (NUT) và đơn vị DPS NAL có DPS NUT. Theo phương án của sáng chế, VPS 804 được chứa trong đơn vị NAL không phải VCL được chỉ định là một đơn vị VPS NAL. Do đó, đơn vị VPS NAL có VPS NUT. Theo phương án của sáng chế, SPS 806 là đơn vị NAL không phải VCL được chỉ định là đơn vị SPS NAL. Do đó, đơn vị SPS NAL có SPS NUT. Theo phương án của sáng chế, PPS 808 được chứa trong đơn vị NAL không phải VCL được chỉ định là đơn vị PPS NAL. Do đó, đơn vị PPS NAL có PPS NUT.

PH 812 là cấu trúc cú pháp mà chứa các phần tử cú pháp mà áp dụng tới tất cả lát (ví dụ, các lát 818) của ảnh được mã hóa (ví dụ, ảnh 814). Theo phương án của sáng chế, PH 812 thuộc loại đơn vị NAL không phải VCL mới được chỉ định là đơn vị PH NAL. Do đó, đơn vị PH NAL có PH NUT (ví dụ, PH_NUT). Theo phương án của sáng chế, có một và chỉ một PH 812 được bao gồm trong mỗi PU 801. Tức là, PU 801 chứa một PH 812 duy nhất. Trong phương án của sáng chế, chính xác một đơn vị PH NAL được hiện diện đối với mỗi ảnh 801 trong dòng bit 800.

Trong một phương án, đơn vị PH NAL được liên kết với PH 812 có ID theo thời gian 817 và ID lớp 813. ID theo thời gian 817 chỉ báo vị trí của đơn vị PH NAL, theo thời điểm, tương ứng với các đơn vị PH NAL khác trong dòng bit (ví dụ, dòng bit 801). ID lớp 813 chỉ báo lớp (ví dụ, lớp 531 hoặc lớp 532) mà chứa đơn vị PH NAL. Trong một phương án, ID theo thời gian 817 tương tự như, nhưng

khác với, POC. POC nhận dạng duy nhất mỗi ảnh theo thứ tự. Trong dòng bit lớp đơn, ID theo thời gian 817 và POC sẽ tương tự nhau. Trong dòng bit đa lớp (ví dụ, xem FIG.5), các ảnh trong cùng một AU có các POC khác nhau, nhưng cùng ID theo thời gian 817.

Trong phương án của sáng chế, đơn vị PH NAL đứng trước đơn vị NAL lớp mã hóa video mà chứa lát thứ nhất 818 của ảnh được kết hợp 814. Điều này thiết lập sự kết hợp giữa PH 812 và các lát 818 của ảnh 814 được kết hợp với PH 812 mà không có sự cần thiết phải có ID thông tin tiêu đề ảnh được báo hiệu trong PH 812 và được tham chiếu từ thông tin tiêu đề lát 820. Do đó, có thể được xem xét rằng tất cả các đơn vị NAL lớp mã hóa video giữa hai PH 812 thuộc về cùng ảnh 814 và ảnh 814 được kết hợp với PH thứ nhất 812 giữa hai PH 812. Trong phương án, đơn vị NAL lớp mã hóa video thứ nhất mà đứng sau PH 812 chứa lát thứ nhất 818 của ảnh 814 (ví dụ, lát từ góc trên cùng bên trái của ảnh) được kết hợp với PH 812 và được chứa trong PU.

Trong phương án của sáng chế, đơn vị NAL thông tin tiêu đề ảnh đứng sau các tập hợp tham số mức ảnh (ví dụ, PPS) hoặc các tập hợp tham số mức cao hơn như DCI (còn gọi là, DPS), VPS, SPS, PPS, v.v mà có cả ID theo thời gian và ID lớp nhỏ hơn ID theo thời gian 817 và ID lớp 813 của đơn vị NAL thông tin tiêu đề ảnh, một cách lần lượt. Do đó, các tập hợp tham số này không được lặp lại trong ảnh hoặc đơn vị truy nhập. Do thứ tự này, PH 812 có thể được phân giải ngay lập tức. Tức là, các tập hợp tham số mà chứa các tham số liên quan đến toàn bộ ảnh được bố trí trong dòng bit trước đơn vị NAL thông tin tiêu đề ảnh. Bất kỳ phần nào mà chứa các tham số đối với một phần của ảnh được bố trí sau đơn vị NAL thông tin tiêu đề ảnh.

Trong một phương án thay thế, đơn vị PH NAL đứng sau các tập hợp tham số mức ảnh và các bản tin thông tin nâng cao bổ sung (SEI-supplemental enhancement information) hoặc các tập hợp tham số cấp cao hơn như DCI (hay còn gọi là DPS), VPS, SPS, PPS, APS, bản tin SEI, v.v.

Trong phương án của sáng chế, PH 812 có thể chứa phần tử cú pháp (ví dụ, cờ, v.v) mà chỉ rõ loại ảnh của ảnh 814 được kết hợp với PH 812. Loại ảnh có thể bao gồm, nhưng không giới hạn, các loại sau đây: Ảnh làm mới bộ giải mã tức thời (IDR-instantaneous decoder refresh), ảnh truy nhập ngẫu nhiên sạch (CRA-clean random access), ảnh GDR, ảnh không phải IRAP mà là ảnh không phải ảnh GDR và chỉ chứa các lát được dự đoán nội bộ (lát I), ảnh không phải IRAP mà là

ảnh không phải GDR và chỉ chứa các lát được dự đoán liên đới đơn hướng (lát P) và lát I, và ảnh không phải IRAP mà là ảnh không phải GDR và chỉ chứa các lát được dự đoán liên đới hai chiều (các lát B), các lát P và các lát I. Do đó, một cờ trong PH 812 có thể chỉ báo rằng tất cả các lát (ví dụ, các lát 818) trong ảnh (ví dụ, ảnh 814) có phải là, ví dụ, các lát của ảnh GDR (ví dụ, ảnh GDR 602) hay không. Điều này cũng hỗ trợ báo hiệu số đếm thứ tự ảnh (POC) điểm khôi phục đối với ảnh GDR một lần trong PH 812 thay vì trong mọi thông tin tiêu đề lát (ví dụ: thông tin tiêu đề lát 820)

Theo phương án của sáng chế, một hoặc nhiều phần tử cú pháp được báo hiệu trong PH 812 thay vì các thông tin tiêu đề lát 820 của ảnh 814 được kết hợp với PH 812. Các phần tử cú pháp này là ký hiệu nhận dạng (ID) PPS được tham chiếu bởi ảnh 814, cờ mà chỉ rõ rằng ảnh 814 có phải là ảnh tham chiếu hay không, mặt phẳng màu của ảnh 814, bit ít quan trọng nhất POC (LSB) của ảnh 814, POC điểm khôi phục nếu ảnh 814 là ảnh GDR (ví dụ, ảnh GDR 602), cờ mà chỉ rõ rằng ảnh trước ảnh 814 có được xuất ra hay không và cờ mà chỉ rõ rằng ảnh có phải là ảnh đầu ra hay không. Ký hiệu nhận dạng (ID) PPS là ký hiệu nhận dạng mà nhận dạng PPS cụ thể đối với ảnh 814. Mặt phẳng màu của ảnh 814 bao gồm các thành phần độ chói và sắc độ (ví dụ: Y, Cb và Cr, v.v.) đối với ảnh 814. POC LSB là bit (hoặc các bit) mà nhận dạng POC. POC là biến được kết hợp với mỗi ảnh (ví dụ, ảnh 814), nhận dạng duy nhất ảnh được kết hợp trong số tất cả các ảnh trong CLVS và, khi ảnh được kết hợp được xuất ra từ DPB, chỉ báo vị trí của ảnh được kết hợp trong thứ tự đầu ra so với các vị trí thứ tự đầu ra của các ảnh khác trong cùng CLVS mà sẽ được xuất ra từ DPB. POC điểm khôi phục là POC mà nhận dạng điểm khôi phục và ảnh điểm khôi phục.

Việc di chuyển các phần tử cú pháp từ các thông tin tiêu đề lát 820 sang PH 812 cho phép các phần tử cú pháp được báo hiệu một lần cho toàn bộ ảnh 814 thay vì lặp lại các phần tử cú pháp trong mỗi thông tin tiêu đề lát 820. Điều này làm giảm dư thừa và tăng hiệu quả mã hóa.

Ảnh 814 là mảng của các mẫu độ chói trong định dạng đơn sắc hoặc mảng các mẫu độ chói và hai mảng mẫu sắc độ tương ứng trong định dạng màu 4:2:0, 4:2:2, và 4:4:4. Trong phương án của sáng chế, có một và chỉ một ảnh 814 được chứa trong mỗi PU 801. Theo đó, có chỉ một PH 812 và chỉ một ảnh 814 tương ứng với PH 812 trong mỗi PU 801. Tức là, PU 801 chứa một hoặc chỉ một ảnh 814.

Ảnh 814 có thể là khung hoặc trường. Tuy nhiên, trong một CVS 816, tất cả ảnh 814 đều là các khung hoặc tất cả ảnh 814 là các trường. CVS 816 là chuỗi video được mã hóa cho mọi chuỗi video lớp được mã hóa (CLVS) trong dòng bit video 800. Đáng chú ý, CVS 816 và CLVS là giống nhau khi dòng bit video 800 bao gồm một lớp. CVS 816 và CLVS chỉ khác nhau khi dòng bit video 800 bao gồm nhiều lớp (ví dụ, như được thể hiện trên FIG.5).

Các PU 801 có thể bao gồm chung CLVS. CLVS là chuỗi của các PU 801 có cùng giá trị `nuh_layer_id` mà bao gồm, theo thứ tự giải mã, (CLVSS) PU bắt đầu chuỗi video lớp được mã hóa, theo sau bởi 0 hoặc nhiều PU 801 không phải là PU CLVSS, bao gồm tất cả các PU tiếp theo 801 tối đa nhưng không bao gồm bất kỳ PU 801 tiếp theo là PU CLVSS. CLVSS PU là PU 801 trong đó ảnh được mã hóa (ví dụ, ảnh 814) là ảnh CLVSS. Ảnh CLVSS là ảnh được mã hóa mà là ảnh IRAP với `NoOutputBeforeRecoveryFlag` bằng 1 hoặc ảnh GDR (ví dụ, ảnh GDR 602) với `NoOutputBeforeRecoveryFlag` bằng 1.

Mỗi ảnh 814 chứa một hoặc nhiều lát 818. Lát 818 là số nguyên của các ô hoàn chỉnh hoặc số nguyên của các hàng CTU hoàn chỉnh liên tiếp trong ô của ảnh (ví dụ, hình 814). Mỗi lát 818 được chứa riêng trong một đơn vị NAL (ví dụ, đơn vị VCL NAL). Theo một phương án, đơn vị NAL đơn được liên kết với hoặc có ID lớp 815. Ô (không được hiển thị) là một vùng hình chữ nhật gồm các CTU trong cột ô cụ thể và hàng ô cụ thể trong ảnh (ví dụ, hình 814). Ô là một phần được phân chia của ảnh được tạo bởi các biên ngang và dọc. Các ô có thể là hình chữ nhật và/hoặc hình vuông. Cụ thể, ô bao gồm bốn cạnh mà được kết nối tại các góc vuông. Bốn cạnh bao gồm hai cặp cạnh song song. Hơn nữa, các cạnh trong cặp cạnh song song có độ dài bằng nhau. Như vậy, ô có thể có dạng chữ nhật bất kỳ, trong đó hình vuông là trường hợp đặc biệt của hình chữ nhật trong đó cả bốn cạnh đều có chiều dài bằng nhau. Ảnh/hình ảnh có thể chứa một hoặc nhiều ô. CTU (không được thể hiện) là CTB của các mẫu độ chói, hai CTB tương ứng của các mẫu sắc độ của ảnh mà có ba mảng mẫu hoặc CTB của các mẫu của ảnh đơn sắc hoặc ảnh mà được mã hóa sử dụng ba mặt phẳng màu riêng biệt và các cấu trúc cú pháp được sử dụng để mã hóa các mẫu. Khối cây mã hóa (CTB) (không được thể hiện) là khối NxN của các mẫu đối với một vài giá trị của N sao cho việc phân chia của thành phần thành các CTB là sự phân chia. Khối (không được thể hiện) là mảng MxN (M-cột x N-hàng) của các mẫu (ví dụ, các điểm ảnh), hoặc mảng MxN của các hệ số biến đổi.

Theo phương án của sáng chế, mỗi lát 818 chứa thông tin tiêu đề lát 820. Thông tin tiêu đề lát 820 là một phần của lát được mã hóa 818 chứa các phần tử dữ liệu liên quan đến tất cả các ô hoặc hàng CTU trong ô được biểu diễn trong lát 818. Tức là là, thông tin tiêu đề lát cắt 820 chứa thông tin về lát 818, chẳng hạn như loại lát, các ảnh tham chiếu nào sẽ được sử dụng, v.v.

Các ảnh 814 và các lát 818 của chúng bao gồm dữ liệu được kết hợp với các ảnh hoặc video được mã hóa hoặc giải mã. Do đó, các ảnh 814 và các lát 818 của chúng có thể được gọi đơn giản là tải trọng hoặc dữ liệu được mang trong dòng bit 800.

Người có hiểu biết trung bình trong lĩnh vực hiểu rằng dòng bit 800 có thể chứa các tham số và thông tin khác trong các ứng dụng thực tế.

FIG.9 là phương án của phương pháp 900 để giải mã được thực hiện bởi bộ giải mã video (ví dụ, bộ giải mã video 400). Phương pháp 900 có thể được thực hiện sau khi dòng bit được thu trực tiếp hoặc gián tiếp từ bộ mã hóa video (ví dụ, bộ mã hóa video 300). Phương pháp 900 cải thiện quá trình giải mã bằng cách đảm bảo rằng PU (ví dụ, PU 801) chứa chỉ một PH (ví dụ, PH 812) và chỉ một ảnh (ví dụ, ảnh 814), mà cho phép ID lớp (ví dụ, ID lớp 813) của đơn vị PH NAL có PH_NUT bằng với ID lớp (ví dụ, ID lớp 815) của đơn vị VCL NAL được liên kết và ID theo thời gian (ví dụ, ID theo thời gian 817) của đơn vị PH NAL bằng với ID theo thời gian (ví dụ, ID theo thời gian 819) của PU mà chứa đơn vị PH NAL. Việc sử dụng ID lớp 813/815 và ID theo thời gian 817/819, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

Trong khối 902, bộ giải mã video thu dòng bit video (ví dụ, dòng bit video 800) bao gồm PU (ví dụ, PU 801). PU chứa đơn vị NAL mà có PH_NUT. PH_NUT biểu thị rằng ID lớp 813 của đơn vị PH NAL bằng ID lớp 815 của đơn vị VCL NAL được liên kết và ID theo thời gian 817 của đơn vị PH NAL bằng ID theo thời gian 819 của PU 801. Trong phương án của sáng chế, đơn vị NAL thông tin tiêu đề ảnh là đơn vị NAL không phải VCL. Trong một phương án, đơn vị PH NAL chứa PH (ví dụ, PH 812).

Trong một phương án, ID lớp 815 của đơn vị VCL NAL nhận dạng lớp (ví dụ, lớp 531, lớp 532, v.v.) mà đơn vị VCL NAL được liên kết thuộc về. Tức là, ID lớp nhận dạng lớp mà bao gồm đơn vị VCL NAL được liên kết. Trong một

phương án, ID lớp nhận dạng lớp của ảnh trong dòng bit video (ví dụ, trong dòng bit video đa lớp). Trong một phương án, ID lớp của một hoặc cả hai đơn vị PH NAL và đơn vị VCL NAL được chỉ định là `nuh_layer_id`. Trong một phương án, đơn vị VCL NAL được liên kết chứa một trong nhiều lát (ví dụ, các lát 818) của ảnh.

Trong một phương án, ID theo thời gian nhận dạng vị trí theo thời điểm của ảnh được mã hóa trong dòng bit video (ví dụ, trong dòng bit video đa lớp). Tức là, ID theo thời gian nhận dạng trong đó ảnh được mã hóa được đặt trong CVS (ví dụ, CVS 816) ứng với các ảnh khác trong CVS về mặt thời gian. Trong một phương án, PU chứa chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh. Việc sử dụng ID lớp 813/815 và ID theo thời gian 817/819, điểm bắt đầu của ảnh trong dòng bit đa lớp có thể được xác định.

Trong khối 904, bộ giải mã video giải mã ảnh được mã hóa từ PU để thu nhận ảnh được giải mã. Trong phương án của sáng chế, việc giải mã PU bao gồm giải mã các lát khác nhau 818 của ảnh 814 theo thông tin hoặc các lệnh trong một hoặc nhiều trong số DCI 802, VPS 804, SPS 806, PPS 808, PH 812, và thông tin tiêu đề lát 820. Sau khi được giải mã, ảnh 814 có thể được sử dụng để tạo ra hoặc sản xuất chuỗi ảnh hoặc video để hiển thị tới người dùng trên màn hình hoặc màn ảnh của thiết bị điện tử (ví dụ, điện thoại thông minh, máy tính bảng, máy tính xách tay, máy tính cá nhân, v.v.).

FIG.10 là phương án của phương pháp 1000 để mã hóa dòng bit video được thực hiện bởi bộ mã hóa video (ví dụ, bộ mã hóa video 300). Phương pháp 1000 có thể được thực hiện khi ảnh (ví dụ, từ video) cần được mã hóa thành dòng bit video và sau đó được truyền tới bộ giải mã video (ví dụ, bộ giải mã video 400). Phương pháp 1000 cải thiện quá trình mã hóa bằng cách đảm bảo rằng PU (ví dụ, PU 801) chứa chỉ một PH (ví dụ, PH 812) và chỉ một ảnh (ví dụ, ảnh 814), mà cho phép ID lớp (ví dụ, ID lớp 813) của đơn vị PH NAL mà có kiểu đơn vị PH NAL (PH_NUT) bằng với ID lớp (ví dụ, ID lớp 815) của đơn vị VCL NAL được liên kết và ID theo thời gian (ví dụ, ID theo thời gian 817) của đơn vị PH NAL bằng với ID theo thời gian (ví dụ, ID theo thời gian 819) của PU mà chứa đơn vị PH NAL. Việc sử dụng ID lớp 813/815 và ID theo thời gian 817/819, điểm bắt đầu của ảnh có thể được xác định trong dòng bit đa lớp mà không gây ra các lỗi mã hóa. Do đó, với vấn đề thực tiễn, hiệu suất của bộ mã hóa-giải mã được cải thiện, mà mang lại trải nghiệm người dùng tốt hơn.

Trong khối 1002, bộ mã hóa video tạo ra PU (ví dụ, PU 801). PU chứa đơn vị NAL mà có PH_NUT. PH_NUT biểu thị rằng ID lớp của đơn vị NAL bằng ID lớp của đơn vị VCL NAL được liên kết và ID theo thời gian của đơn vị NAL bằng ID theo thời gian của PU. Theo phương án của sáng chế, đơn vị NAL là đơn vị NAL không phải VCL. Trong phương án của sáng chế, đơn vị NAL chứa PH (ví dụ, PH 812).

Trong một phương án, ID lớp 815 của đơn vị VCL NAL nhận dạng lớp (ví dụ, lớp 531, lớp 532, v.v.) mà đơn vị VCL NAL được liên kết thuộc về. Tức là, ID lớp 815 nhận dạng lớp mà bao gồm đơn vị VCL NAL được liên kết. Trong một phương án, ID lớp 813 của đơn vị PH NAL nhận dạng lớp của ảnh trong dòng bit video (ví dụ, trong dòng bit video đa lớp). Trong một phương án, ID lớp được chỉ định là `nuh_layer_id`. Trong một phương án, đơn vị VCL NAL được liên kết chứa một trong nhiều lát (ví dụ, các lát 818) của ảnh.

Trong một phương án, ID theo thời gian của đơn vị PH NAL nhận dạng vị trí theo thời điểm của ảnh được mã hóa trong dòng bit video (ví dụ, trong dòng bit video đa lớp). Tức là, ID theo thời gian của đơn vị PH NAL nhận dạng trong đó ảnh được mã hóa được đặt trong CVS (ví dụ, CVS 816) ứng với các ảnh khác trong CVS về mặt thời gian. Trong một phương án, PU chứa chỉ một thông tin tiêu đề ảnh (PH) và chỉ một ảnh. Nhờ sử dụng ID lớp và ID theo thời gian, điểm bắt đầu của ảnh trong dòng bit đa lớp có thể được xác định.

Trong khối 1004, bộ mã hóa video mã hóa PU thành dòng bit video (ví dụ, dòng bit 800). Trong khối 1006, bộ mã hóa video lưu trữ dòng bit video để truyền thông tới bộ giải mã video. Dòng bit video có thể được lưu trữ trong bộ nhớ cho đến khi dòng bit video được truyền tới bộ giải mã video. Sau khi được nhận bởi bộ giải mã video, dòng bit video được mã hóa có thể được giải mã (ví dụ, như được mô tả nêu trên) để tạo ra hoặc sinh ra chuỗi ảnh hoặc video để hiển thị tới người dùng trên màn hình hoặc màn ảnh của thiết bị điện tử (ví dụ, điện thoại thông minh, máy tính bảng, máy tính xách tay, máy tính cá nhân, v.v.)

Cú pháp và ngữ nghĩa sau có thể được sử dụng để thực hiện các phương án được bộc lộ ở đây. Phần mô tả sau đây là liên quan đến văn bản cơ sở, là tiêu chuẩn kỹ thuật dự thảo VVC mới nhất. Nói cách khác, chỉ đen-ta được mô tả, trong khi văn bản trong văn bản cơ sở không được đề cập dưới đây được áp dụng. Văn bản được bổ sung liên quan đến văn bản cơ sở được in đậm và văn bản bị loại bỏ được thể hiện ở dạng in nghiêng.

Cú pháp RBSP của thông tin tiêu đề ảnh.

picture header_rbsp() {	Nhãn mô tả
pic_type	u(3)
pic_parameter_set_id	ue(v)
non_reference_picture_flag	u(1)
if(separate_colour_plane_flag == 1)	
colour_plane_id	u(2)
pic_order_cnt_lsb	u(v)
if(pic_type == 2)	
recovery_poc_cnt	ue(v)
if(pic_type == 0 pic_type == 1 pic_type == 2)	
no_output_of_prior_pics_flag	u(1)
if(output_flag_present_flag)	
pic_output_flag	u(1)
if (pic_type != 0 && pic_type != 1 && pic_type != 3)	
if(sps_temporal_mvp_enabled_flag && !pps_temporal_mvp_enabled_idc)	
pic_temporal_mvp_enabled_flag	u(1)
rbbsp_trailing_bits()	
}	

Cú pháp RBSP của thông tin tiêu đề lát.

	Nhãn mô tả
<code>slice_header() {</code>	
<code> slice_pic_parameter_set_id</code>	<code>ue(v)</code>
<code> if(rect_slice_flag NumBricksInPic > 1)</code>	
<code> slice_address</code>	<code>u(v)</code>
<code> if(!rect_slice_flag && !single_brick_per_slice_flag)</code>	
<code> num_bricks_in_slice_minus1</code>	<code>ue(v)</code>
<code> non_reference_picture_flag</code>	<code>u(1)</code>
<code> if(pic_type != 0 && pic_type != 1 && pic_type != 3)</code>	
<code> slice_type</code>	<code>ue(v)</code>
<code> if(separate_colour_plane_flag == 1)</code>	
<code> colour_plane_id</code>	<code>u(2)</code>
<code> slice_pic_order_cnt_lsb</code>	<code>u(v)</code>
<code> if(nal_unit_type == GDR_NUT)</code>	
<code> recovery_poc_cnt</code>	<code>ue(v)</code>
<code> if(nal_unit_type == IDR_W_RADL nal_unit_type == IDR_N_LP </code> <code> nal_unit_type == CRA_NUT nal_unit_type == GDR_NUT)</code>	
<code> no_output_of_prior_pics_flag</code>	<code>u(1)</code>
<code> if(output_flag_present_flag)</code>	
<code> pic_output_flag</code>	<code>u(1)</code>
<code> ...</code>	
<code> if(slice_type != I) {</code>	
<code> if(sps_temporal_mvp_enabled_flag &&</code> <code> !pps_temporal_mvp_enabled_idc)</code>	
<code> slice_temporal_mvp_enabled_flag</code>	<code>u(1)</code>
<code> if(slice_type == B && !pps_mvd_l1_zero_idc)</code>	
<code> mvd_l1_zero_flag</code>	<code>u(1)</code>
<code> ...</code>	
<code> }</code>	
<code> ...</code>	
<code>}</code>	

Ngữ nghĩa RBSP thông tin tiêu đề ảnh.

Thông tin tiêu đề ảnh chứa thông tin chung cho tất cả các lát của ảnh được mã hóa mà đơn vị VCL NAL tiếp theo trong thứ tự giải mã là lát được mã hóa thứ nhất.

`pic_type` chỉ báo đặc trưng của các ảnh được mã hóa như liệt kê trong Bảng 7-3 đối với giá trị định trước của `pic_type`. Giá trị của `pic_type` sẽ bằng 0 đến 5, trong các dòng bit mà tuân theo phiên bản của tiêu chuẩn kỹ thuật này. Các giá trị khác của `pic_type` được dành riêng cho việc sử dụng tương lai bởi ITU-T | ISO/IEC. Các bộ giải mã mà tuân theo phiên bản này của tiêu chuẩn kỹ thuật này sẽ bỏ qua các giá trị dành riêng của `pic_type`.

Bảng 7-3 - Diễn giải của tham số pic_type

pic_type	đặc trưng của ảnh được mã hóa
0	Ảnh được mã hóa là ảnh IDR.
1	Ảnh được mã hóa là ảnh CRA.
2	Ảnh được mã hóa là ảnh GDR.
3	Ảnh được mã hóa là ảnh không phải IRAP, không phải GDR và chứa chỉ các lát I.
4	Ảnh được mã hóa là ảnh không IRAP, không phải GDR và chỉ chứa các lát P và I.
5	Ảnh được mã hóa là ảnh không phải IRAP, không phải là GDR và chứa các lát cắt B, P và I.

pic_parameter_set_id chỉ rõ giá trị của pps_pic_parameter_set_id đối với PPS được sử dụng. Giá trị của pic_parameter_set_id sẽ nằm trong đoạn từ 0 đến 63.

non_reference_picture_flag bằng 1 chỉ rõ ảnh được kết hợp với thông tin tiêu đề ảnh không bao giờ được sử dụng như là ảnh tham chiếu. non_reference_picture_flag bằng 0 chỉ rõ ảnh có thể có hoặc có thể không được sử dụng như là ảnh tham chiếu.

colour_plane_id chỉ rõ mặt phẳng màu được kết hợp với ảnh được kết hợp với thông tin tiêu đề ảnh khi separate_colour_plane_flag bằng 1. Giá trị của colour_plane_id sẽ nằm trong đoạn từ 0 đến 2. các giá trị colour_plane_id 0, 1 và 2 tương ứng với các mặt phẳng Y, Cb và Cr, một cách lần lượt.

LƯU Ý -: Không có sự phụ thuộc giữa các xử lý giải mã của các ảnh mà có các giá trị colour_plane_id khác nhau .

pic_order_cnt_lsb chỉ rõ môđun số đếm thứ tự ảnh MaxPicOrderCntLsb đối với ảnh được kết hợp với thông tin tiêu đề ảnh. Độ dài của phần tử cú pháp pic_order_cnt_lsb là $\log_2_max_pic_order_cnt_lsb_minus4 + 4$ bit. Giá trị của pic_order_cnt_lsb sẽ nằm trong đoạn từ 0 đến MaxPicOrderCntLsb - 1.

recovery_poc_cnt chỉ rõ điểm khôi phục của các ảnh được giải mã trong thứ tự đầu ra. Nếu có ảnh picA mà đứng sau ảnh GDR hiện tại trong thứ tự giải mã trong CVS và có PicOrderCntVal bằng PicOrderCntVal của ảnh GDR hiện tại cộng giá trị của recovery_poc_cnt, ảnh picA được gọi là ảnh điểm khôi phục. Nếu không phải, ảnh thứ nhất trong thứ tự đầu ra mà có

PicOrderCntVal lớn hơn PicOrderCntVal của ảnh hiện tại cộng giá trị của recovery_poc_cnt được gọi là ảnh điểm khôi phục. Ảnh điểm khôi phục sẽ không đứng trước ảnh GDR hiện tại trong thứ tự giải mã. Giá trị của recovery_poc_cnt sẽ nằm trong đoạn từ 0 đến MaxPicOrderCntLsb – 1.

Biến RpPicOrderCntVal được thu nhận như sau:

$$\text{RpPicOrderCntVal} = \text{PicOrderCntVal} + \text{recovery_poc_cnt}$$

(7-94)

no_output_of_prior_pics_flag ảnh hưởng đến đầu ra của các ảnh được giải mã trước đó trong bộ đệm ảnh được giải mã sau việc giải mã ảnh CLVSS mà không phải là ảnh thứ nhất trong dòng bit như quy định trong Phụ lục C.

pic_output_flag ảnh hưởng tới đầu ra ảnh được giải mã và các xử lý loại bỏ như được quy định trong Phụ lục C. Khi pic_output_flag không được hiện diện, tham số này được xem là bằng 1.

pic_temporal_mvp_enabled_flag chỉ rõ rằng các biến dự đoán vector chuyển động thời gian có thể được sử dụng hay không cho việc dự đoán liên đới. Nếu pic_temporal_mvp_enabled_flag bằng 0, các phần tử cú pháp của ảnh được kết hợp với thông tin tiêu đề ảnh sẽ bị giới hạn sao cho không có biến dự đoán vector chuyển động thời gian được sử dụng trong việc giải mã ảnh. Nếu không phải (pic_temporal_mvp_enabled_flag bằng 1), các biến dự đoán vector chuyển động thời gian có thể được sử dụng trong việc giải mã ảnh.

Khi pic_temporal_mvp_enabled_flag không được hiện diện, những điều sau sẽ áp dụng:

- Nếu sps_temporal_mvp_enabled_flag bằng 0, giá trị của pic_temporal_mvp_enabled_flag được xem là bằng 0.
- Nếu không phải (sps_temporal_mvp_enabled_flag bằng 1), giá trị của pic_temporal_mvp_enabled_flag được xem là bằng pps_temporal_mvp_enabled_idc – 1.

Ngữ nghĩa thông tin tiêu đề đơn vị NAL.

nuh_layer_id chỉ rõ ký hiệu nhận dạng của lớp mà đơn vị VCL NAL thuộc về hoặc ký hiệu nhận dạng của lớp mà đơn vị NAL không phải VCL áp dụng tới.

Giá trị của `nuh_layer_id` là giống nhau đối với tất cả các đơn vị NAL lớp mã hóa video của ảnh được mã hóa. Giá trị của `nah_layer_id` của ảnh được mã hóa hoặc đơn vị truy nhập lớp là giá trị của `nah_layer_id` của các đơn vị NAL lớp mã hóa video của ảnh được mã hóa hoặc đơn vị truy nhập lớp.

Khi `nal_unit_type` bằng `PH_NUT`, giá trị của `nuh_layer_id` sẽ bằng giá trị của `nuh_layer_id` của các đơn vị NAL lớp mã hóa video của đơn vị truy nhập lớp mà chứa đơn vị NAL thông tin tiêu đề ảnh.

Bảng 7-1 - Các mã loại đơn vị NAL và các lớp loại đơn vị NAL

<code>nal_unit_type</code>	Tên của <code>nal_unit_type</code>	Nội dung của đơn vị NAL và cấu trúc cú pháp RBSP	Lớp loại đơn vị NAL
0	<code>TRAIL_NUT</code>	Lát được mã hóa của ảnh kế tiếp <code>slice_layer_rbsp()</code>	VCL
1	<code>STSA_NUT</code>	Lát được mã hóa của ảnh STSA <code>slice_layer_rbsp()</code>	VCL
2	<code>RASL_NUT</code>	Lát được mã hóa của ảnh RASL <code>slice_layer_rbsp()</code>	VCL
3	<code>RADL_NUT</code>	Lát được mã hóa của ảnh RADL <code>slice_layer_rbsp()</code>	VCL
4..7	<code>RSV_VCL_4..</code> <code>RSV_VCL_7</code>	Các loại đơn vị VCL NAL không IRAP được dành riêng	VCL
8 9	<code>IDR_W_RADL</code> <code>IDR_N_LP</code>	Lát mã hóa của ảnh IDR <code>slice_layer_rbsp()</code>	VCL
10	<code>CRA_NUT</code>	Lát được mã hóa của ảnh CRA <code>silce_layer_rbsp()</code>	VCL
11	<code>GDR_NUT</code>	Lát được mã hóa của ảnh GDR <code>slice_layer_rbsp()</code>	VCL
12 13	<code>RSV_IRAP_VCL12</code> <code>RSV_IRAP_VCL13</code>	Các loại đơn vị VCL NAL IRAP được dành riêng	VCL
14..15	<code>RSV_VCL14..</code> <code>RSV_VCL15</code>	Các loại đơn vị VCL NAL không IRAP được dành riêng	VCL
16	<code>SPS_NUT</code>	Tập hợp tham số chuỗi <code>seq_parameter_set_rbsp()</code>	không phải VCL
17	<code>PPS_NUT</code>	Tập hợp tham số ảnh <code>pic_parameter_set_rbsp()</code>	không phải VCL
18	<code>APS_NUT</code>	Tập hợp tham số thích ứng <code>adaptation_parameter_set_rbsp()</code>	không phải VCL

19	PH_NUT	Thông tin tiêu đề ảnh picture_header_rbsp()	không phải VCL
20	EOS_NUT	Kết thúc của chuỗi end_of_seq_rbsp()	không phải VCL
21	EOB_NUT	Kết thúc của dòng bit end_of_bitstream_rbsp()	không phải VCL
22, 23	PREFIX_SEI_NUT SUFFIX_SEI_NUT	Thông tin nâng cao bổ sung sei_rbsp()	không phải VCL
24	DPS_NUT	Tập hợp tham số giải mã decoding_parameter_set_rbsp()	không phải VCL
25..27	RSV_NVCL25 .. RSV_NVCL27	Các loại đơn vị NAL không phải VCL được dành riêng	không phải VCL
28..31	UNSPEC28.. UNSPEC31	Các loại đơn vị NAL không phải VCL không được xác định	không phải VCL

nuh_temporal_id_plus1 trừ 1 chỉ rõ ký hiệu nhận dạng thời gian đối với đơn vị NAL.

Giá trị của nuh_temporal_id_plus1 sẽ không bằng 0.

Biến TemporalId được thu nhận như sau:

$$\text{TemporalId} = \text{nuh_temporal_id_plus1} - 1 \quad (7-1)$$

Khi nal_unit_type nằm trong đoạn từ IDR_W_RADL đến RSV_IRAP_VCL13, TemporalId sẽ bằng 0.

Khi nal_unit_type bằng STSA_NUT, TemporalId sẽ không bằng 0.

Giá trị của TemporalId sẽ giống nhau đối với tất cả các đơn vị NAL lớp mã hóa video của đơn vị truy nhập lớp. Giá trị của TemporalId of ảnh được mã hóa hoặc đơn vị truy nhập lớp là giá trị của TemporalId của các đơn vị NAL lớp mã hóa video của ảnh được mã hóa hoặc đơn vị truy nhập lớp. [Ed. (YK): Kiểm tra xem tất cả các AU lớp trong AU có được yêu cầu có cùng giá trị của TemporalId hay không.] Giá trị của TemporalId của đại diện lớp con là giá trị lớn nhất của TemporalId của tất cả các đơn vị NAL lớp mã hóa video trong đại diện lớp con.

Giá trị của TemporalId đối với các đơn vị NAL không phải lớp mã hóa video bị giới hạn như sau:

– Nếu `nal_unit_type` bằng `DPS_NUT`, `VPS_NUT`, hoặc `SPS_NUT`, `TemporalId` bằng 0 và `TemporalId` của đơn vị truy nhập lớp mà chứa đơn vị NAL sẽ bằng 0.

– Nếu không phải, nếu `nal_unit_type` bằng `PH_NUT`, **`TemporalId` sẽ bằng `TemporalId` của đơn vị truy nhập lớp mà chứa đơn vị NAL.**

– Nếu không phải, khi `nal_unit_type` không bằng `EOS_NUT` và không bằng `EOB_NUT`, `TemporalId` sẽ lớn hơn hoặc bằng `TemporalId` của đơn vị truy nhập lớp mà chứa đơn vị NAL.

Lưu ý 5 – khi đơn vị NAL là đơn vị NAL không phải VCL, giá trị của `TemporalId` bằng giá trị nhỏ nhất của các giá trị `TemporalId` của tất cả đơn vị truy nhập lớp mà đơn vị NAL không phải VCL áp dụng tới. Khi `nal_unit_type` bằng `PPS_NUT` hoặc `APS_NUT`, `TemporalId` có thể lớn hơn hoặc bằng `TemporalId` của đơn vị truy nhập lớp chứa, khi tất cả PPS và APS có thể được chứa trong phần bắt đầu của dòng bit, trong đó ảnh được mã hóa thứ nhất có `TemporalId` bằng 0. Khi `nal_unit_type` bằng `PREFIX_SEI_NUT` hoặc `SUFFIX_SEI_NUT`, `TemporalId` có thể lớn hơn hoặc bằng `TemporalId` của đơn vị truy nhập lớp chứa, khi đơn vị SEI NAL có thể chứa thông tin mà áp dụng tới tập con dòng bit mà bao gồm các đơn vị truy nhập lớp mà các giá trị `TemporalId` lớn hơn `TemporalId` của đơn vị truy nhập lớp mà chứa đơn vị SEI NAL.

Thứ tự của các đơn vị NAL và các ảnh được mã hóa và sự kết hợp của chúng với các đơn vị truy nhập lớp và các đơn vị truy nhập.

Phần này quy định thứ tự của các đơn vị NAL và ảnh được mã hóa và sự kết hợp của chúng với các đơn vị truy nhập lớp và các đơn vị truy nhập đối với các CVS mà tuân theo một hoặc nhiều cấu hình được chỉ định trong Phụ lục A và được giải mã nhờ sử dụng xử lý giải mã được quy định trong các điều từ 2 đến 10.

Đơn vị truy nhập lớp bao gồm **một đơn vị NAL thông tin tiêu đề ảnh**, một ảnh được mã hóa, **mà bao gồm một hoặc nhiều đơn vị NAL lớp mã hóa video**, và 0 hoặc nhiều hơn các đơn vị NAL lớp mã hóa video và 0 hoặc nhiều hơn các đơn vị NAL không phải lớp mã hóa video. Sự kết hợp của các đơn vị VCL NAL với các ảnh được mã hóa được mô tả trong điều 7.4.2.4.4.

Đơn vị truy nhập bao gồm đơn vị NAL dấu phân cách đơn vị truy nhập và một hoặc nhiều đơn vị truy nhập lớp theo thứ tự tăng dần của `nuh_layer_id`.

Đơn vị truy nhập thứ nhất trong dòng bit bắt đầu bằng đơn vị NAL thứ nhất của dòng bit.

Đặt **firstPicHeaderNalUnitInAu** là đơn vị NAL thông tin tiêu đề ảnh là thông tin tiêu đề ảnh của ảnh được mã hóa thứ nhất mà **PicOrderCntVal** được thu nhận khác với **PicOrderCntVal** của ảnh được mã hóa trước đó. Đơn vị đầu tiên trong số bất kỳ đơn vị NAL mà đứng trước **firstPicHeaderNalUnitInAu** và kế tiếp đơn vị VCL NAL cuối cùng mà đứng trước **firstPicHeaderNalUnitInAu**, nếu có, chỉ rõ sự bắt đầu của đơn vị truy nhập mới:

- Đơn vị DPS NAL (khi được hiện diện),
- Đơn vị VPS NAL (khi được hiện diện),
- Đơn vị SPS NAL (khi được hiện diện),
- Đơn vị PPS NAL (khi được hiện diện),
- Đơn vị APS NAL (khi được hiện diện),
- Đơn vị SEI NAL tiền tố (khi được hiện diện),
- các đơn vị NAL với **nal_unit_type** trong phạm vi của **RSV_NVCL_25..RSV_NVCL_26** (khi được hiện diện),
- các đơn vị NAL với **nal_unit_type** trong phạm vi của **UNSPEC28..UNSPEC29** (khi được hiện diện).

LƯU Ý - Đơn vị NAL thứ nhất mà đứng trước **firstPicHeaderNalUnitInAu** và kế tiếp đơn vị VCL NAL cuối cùng mà đứng trước **firstPicHeaderNalUnitInAu**, nếu có, chỉ có thể là một trong các đơn vị NAL được liệt kê nêu trên.

Khi không có đơn vị NAL nêu trên mà đứng trước **firstPicHeaderNalUnitInAu** và kế tiếp VCL NAL cuối cùng mà đứng trước **firstPicHeaderNalUnitInAu**, nếu có, **firstPicHeaderNalUnitInAu** bắt đầu đơn vị truy nhập mới.

Thứ tự của các ảnh được mã hóa và các đơn vị NAL không phải VCL trong đơn vị truy nhập lớp hoặc đơn vị truy nhập phải tuân theo các điều kiện sau đây:

– Mỗi đơn vị truy nhập lớp sẽ bao gồm một và chỉ một đơn vị NAL thông tin tiêu đề ảnh, mà sẽ đứng trước đơn vị NAL lớp mã hóa video thứ nhất của đơn vị truy nhập lớp.

– Khi bất kỳ các đơn vị DPS NAL, các đơn vị VPS NAL, các đơn vị SPS NAL, hoặc các đơn vị PPS NAL được hiện diện trong đơn vị truy nhập lớp, chúng sẽ đứng trước đơn vị NAL thông tin tiêu đề ảnh của đơn vị truy nhập lớp.

- Khi bất kỳ các đơn vị APS NAL, đơn vị SEI NAL tiền tố, đơn vị NAL với `nal_unit_type` trong phạm vi `RSV_NVCL_25..RSV_NVCL_26` hoặc các đơn vị NAL với `nal_unit_type` trong phạm vi `UNSPEC28..UNSPEC29` được hiện diện trong đơn vị truy nhập lớp, chúng sẽ không đứng sau đơn vị VCL NAL cuối cùng của đơn vị truy nhập lớp.

- Các đơn vị NAL mà có `nal_unit_type` bằng `SUFFIX_SEI_NUT` hoặc `RSV_NVCL_27` hoặc trong phạm vi `UNSPEC30..UNSPEC31` trong đơn vị truy nhập lớp sẽ không đứng trước đơn vị VCL NAL thứ nhất của đơn vị truy nhập lớp.

- Khi kết thúc của đơn vị NAL chuỗi được hiện diện trong đơn vị truy nhập, nó sẽ là đơn vị NAL cuối cùng trong số tất cả các đơn vị NAL trong đơn vị truy nhập khác với kết thúc của đơn vị NAL dòng bit (khi được hiện diện).

– Khi kết thúc của đơn vị NAL dòng bit được hiện diện trong đơn vị truy nhập, nó sẽ là đơn vị NAL cuối cùng trong đơn vị truy nhập.

Ngữ nghĩa RBSP thông tin tiêu đề lát.

Khi được hiện diện, giá trị của mỗi phần tử cú pháp thông tin tiêu đề lát `slice_pic_parameter_set_id`, `non_reference_picture_flag`, `colour_plane_id`, `slice_pic_order_cnt_lsb`, `recovery_poc_cnt`, `no_output_of_prior_pics_flag`, `pic_output_flag`, và `slice_temporal_mvp_enabled_flag` sẽ giống nhau trong tất cả các thông tin tiêu đề lát của ảnh được mã hóa.

Biến `CuQpDeltaVal`, chỉ rõ độ chênh lệch giữa tham số lượng tử hóa độ chói đối với đơn vị mã hóa mà chứa `cu_qp_delta_abs` và dự đoán của nó, được thiết lập bằng 0. Các biến `CuQpOffsetCb`, `CuQpOffsetCr`, và `CuQpOffsetCbCr`, chỉ định các giá trị cần được sử dụng khi xác định các giá trị tương ứng của các tham số lượng tử hóa Qp'_{Cb} , Qp'_{Cr} , và Qp'_{CbCr} cho đơn vị mã hóa mà chứa `cu_chroma_qp_offset_flag`, đều được đặt bằng 0.

slice_pic_parameter_set_id chỉ rõ giá trị của pps_pic_parameter_set_id đối với PPS được sử dụng. Giá trị của slice_pic_parameter_set_id sẽ nằm trong đoạn từ 0 đến 63.

Đây là yêu cầu của sự tương thích dòng bit mà giá trị của TemporalId của ảnh hiện tại lớn hơn hoặc bằng giá trị của TemporalId của PPS mà có pps_pic_parameter_set_id bằng slice_pic_parameter_set_id

slice_address chỉ rõ địa chỉ lát của lát. Khi không được hiện diện, giá trị của slice_address được xem là bằng 0.

Nếu rect_slice_flag bằng 0, phần sau đây sẽ áp dụng:

- Địa chỉ lát là ID viên như được chỉ rõ bởi công thức (7-59).
- Độ dài của slice_address là $\text{Ceil}(\text{Log2}(\text{NumBricksInPic}))$ bit.
- Giá trị của slice_address sẽ nằm trong đoạn từ 0 đến $\text{NumBricksInPic} - 1$.

Nếu không phải (rect_slice_flag bằng 1), phần sau đây được áp dụng:

- Địa chỉ lát là ID lát của lát.
- Độ dài của slice_address là $\text{signalled_slice_id_length_minus1} + 1$ bit.

Nếu signalled_slice_id_flag bằng 0, giá trị của slice_address sẽ nằm trong đoạn từ 0 đến $\text{num_slices_in_pic_minus1}$. Nếu không phải, giá trị của slice_address sẽ nằm trong đoạn từ 0 đến $2^{(\text{signalled_slice_id_length_minus1} + 1)} - 1$.

Đây là yêu cầu về sự tương thích dòng bit mà các điều kiện sau đây được áp dụng:

- Giá trị của slice_address sẽ không bằng giá trị của slice_address của bất kỳ đơn vị NAL lát được mã hóa khác của cùng ảnh được mã hóa.
- Khi rect_slice_flag bằng 0, các lát của ảnh sẽ nằm trong thứ tự tăng dần của các giá trị slice_address của chúng.
- Dạng của các lát của ảnh là phải sao cho mỗi viên, khi được giải mã, sẽ có toàn bộ biên bên trái và toàn bộ biên trên cùng bao gồm biên ảnh hoặc bao gồm các biên của (các) viên đã được giải mã trước đó.

num_bricks_in_slice_minus1, khi được hiện diện, chỉ rõ số lượng biên trong lát trừ 1. Giá trị của num_bricks_in_slice_minus1 sẽ nằm trong đoạn từ 0 đến

NumBricksInPic – 1. Khi rect_slice_flag bằng 0 và single_brick_per_slice_flag bằng 1, giá trị của num_bricks_in_slice_minus1 được xem là bằng 0. Khi single_brick_per_slice_flag bằng 1, giá trị của num_bricks_in_slice_minus1 được xem là bằng 0.

Biến NumBricksInCurrSlice, mà chỉ rõ số lượng viên trong lát hiện tại, và SliceBrickIdx[i], mà chỉ rõ chỉ số viên của viên thứ i trong lát hiện tại, được thu nhận như sau:

```

if( rect_slice_flag ) {
    sliceIdx = 0
    while( slice_address != slice_id[ sliceIdx ] )
        sliceIdx++
    NumBricksInCurrSlice = NumBricksInSlice[ sliceIdx ]
    brickIdx = TopLeftBrickIdx[ sliceIdx ]
    for( bIdx = 0; brickIdx <= BottomRightBrickIdx[ sliceIdx ]; brickIdx++ )
        (7-92)
        if( BricksToSliceMap[ brickIdx ] == sliceIdx )
            SliceBrickIdx[ bIdx++ ] = brickIdx
        } else {
    NumBricksInCurrSlice = num_bricks_in_slice_minus1 + 1
    SliceBrickIdx[ 0 ] = slice_address
    for( i = 1; i < NumBricksInCurrSlice; i++ )
        SliceBrickIdx[ i ] = SliceBrickIdx[ i - 1 ] + 1
    }

```

Các biến SubPicIdx, SubPicLeftBoundaryPos, SubPicTopBoundaryPos, SubPicRightBoundaryPos, và SubPicBotBoundaryPos được thu nhận như sau:

```

SubPicIdx =
    CtbToSubPicIdx[ CtbAddrBsToRs[ FirstCtbAddrBs[ SliceBrickIdx[ 0 ] ] ]
    ]
if( subpic_treated_as_pic_flag[ SubPicIdx ] ) {
    SubPicLeftBoundaryPos =
        SubPicLeft[ SubPicIdx ] * ( subpic_grid_col_width_minus1 + 1 ) * 4
    SubPicRightBoundaryPos =
        ( SubPicLeft[ SubPicIdx ] + SubPicWidth[ SubPicIdx ] ) *

```

$$\begin{aligned}
 & (\text{subpic_grid_col_width_minus1} + 1) * 4 & (7-93) \\
 \text{SubPicTopBoundaryPos} = & \\
 \text{SubPicTop[SubPicIdx]} * (\text{subpic_grid_row_height_minus1} + 1) * 4 & \\
 \text{SubPicBotBoundaryPos} = & \\
 (\text{SubPicTop[SubPicIdx]} + \text{SubPicHeight[SubPicIdx]}) * & \\
 (\text{subpic_grid_row_height_minus1} + 1) * 4 & \\
 \} &
 \end{aligned}$$

non_reference_picture_flag bằng 1 chỉ rõ rằng ảnh mà chứa lát không bao giờ được sử dụng như là ảnh tham chiếu. *non_reference_picture_flag* bằng 0 chỉ rõ rằng ảnh mà chứa lát có thể có hoặc có thể không được sử dụng như là ảnh tham chiếu.

slice_type chỉ rõ loại mã hóa của lát theo Bảng 7-11.

Bảng 7-11 - Kết hợp tên gọi với *Slice_type*

<i>slice_type</i>	Tên của <i>slice_type</i>
0	B (lát B)
1	P (lát P)
2	I (lát I)

Khi *nal_unit_type* là giá trị của *nal_unit_type* trong đoạn từ *IDR_W_RADL* đến *CRA_NUT*, và ảnh hiện tại là ảnh thứ nhất trong đơn vị truy nhập, *slice_type* sẽ bằng 2. **Khi không được hiện diện, giá trị của *slice_type* được xem là bằng 2.**

colour_plane_id chỉ rõ mặt phẳng màu được kết hợp với RBSP lát hiện tại khi *separate_colour_plane_flag* bằng 1. Giá trị của *colour_plane_id* sẽ nằm trong đoạn từ 0 đến 2. các giá trị *colour_plane_id* 0, 1 và 2 tương ứng với các mặt phẳng Y, Cb và Cr, một cách lần lượt.

LƯU Ý -: Không có sự phụ thuộc giữa các xử lý giải mã của các ảnh mà có các giá trị *colour_plane_id* khác nhau.

slice_pic_order_cnt_lsb chỉ rõ môđun số đếm thứ tự ảnh *MaxPicOrderCntLsb* đối với ảnh hiện tại. Độ dài của phần tử cú pháp *slice_pic_order_cnt_lsb* là $\log_2 \text{max_pic_order_cnt_lsb_minus4} + 4$ bit. Giá trị của *slice_pic_order_cnt_lsb* sẽ nằm trong đoạn từ 0 đến *MaxPicOrderCntLsb* - 1.

recovery_poc_cnt chỉ rõ điểm khôi phục của các ảnh được giải mã trong thứ tự đầu ra. Nếu có ảnh *picA* mà đứng sau ảnh GDR hiện tại trong thứ tự giải mã trong CVS và có *PicOrderCntVal* bằng *PicOrderCntVal* của ảnh GDR hiện tại cộng giá trị của *recovery_poc_cnt*, ảnh *picA* được gọi là ảnh điểm khôi phục. Nếu không phải, ảnh thứ nhất trong thứ tự đầu ra mà có *PicOrderCntVal* lớn hơn *PicOrderCntVal* của ảnh hiện tại cộng giá trị của *recovery_poc_cnt* được gọi là ảnh điểm khôi phục. Ảnh điểm khôi phục sẽ không đứng trước ảnh GDR hiện tại trong thứ tự giải mã. Giá trị của *recovery_poc_cnt* sẽ nằm trong đoạn từ 0 đến $\text{MaxPicOrderCntLsb} - 1$.

Biến *RpPicOrderCntVal* được thu nhận như sau:

$$\text{RpPicOrderCntVal} = \text{PicOrderCntVal} + \text{recovery_poc_cnt} \quad (7-94)$$

no_output_of_prior_pics_flag ảnh hưởng đến đầu ra của các ảnh được giải mã trước đó trong bộ đệm ảnh được giải mã sau việc giải mã ảnh CLVSS mà không phải là ảnh thứ nhất trong dòng bit như quy định trong Phụ lục C.

pic_output_flag ảnh hưởng tới đầu ra ảnh được giải mã và các xử lý loại bỏ như được quy định trong Phụ lục C. Khi *pic_output_flag* không được hiện diện, tham số này được xem là bằng 1.

lice_temporal_mvp_enabled_flag chỉ rõ rằng các biến dự đoán vector chuyển động thời gian có thể được sử dụng hay không cho việc dự đoán liên đới. Nếu *slice_temporal_mvp_enabled_flag* bằng 0, các phần tử cú pháp của ảnh hiện tại sẽ bị giới hạn sao cho không có biến dự đoán vector chuyển động thời gian được sử dụng trong việc giải mã ảnh hiện tại. Nếu không phải (*slice_temporal_mvp_enabled_flag* bằng 1), các biến dự đoán vector chuyển động thời gian có thể được sử dụng trong việc giải mã ảnh hiện tại.

Khi *slice_temporal_mvp_enabled_flag* không được hiện diện, phần sau đây được áp dụng:

Nếu *sps_temporal_mvp_enabled_flag* bằng 0, giá trị của *slice_temporal_mvp_enabled_flag* được xem là bằng 0.

Nếu không phải (*sps_temporal_mvp_enabled_flag* bằng 1), giá trị của *slice_temporal_mvp_enabled_flag* được xem là bằng $\text{pps_temporal_mvp_enabled_idc} - 1$.

mvd_l1_zero_flag bằng 1 chỉ báo rằng cấu trúc cú pháp *mvd_coding*(*x0*, *y0*, 1) không được phân tích và *MvdL1*[*x0*][*y0*][*compIdx*] và *MvdL1*[*x0*][*y0*][*cpIdx*][*compIdx*] được thiết lập bằng 0 đối với *compIdx* = 0..1 và *cpIdx* = 0..2. *mvd_l1_zero_flag* bằng 0 chỉ báo rằng cấu trúc cú pháp *mvd_coding*(*x0*, *y0*, 1) được phân tích. Khi không được hiện diện, giá trị của *mvd_l1_zero_flag* được xem là bằng *pps_mvd_l1_zero_idc* - 1.

Xử lý giải mã đối với số đếm thứ tự ảnh.

Đầu ra của xử lý này là *PicOrderCntVal*, số đếm thứ tự ảnh của ảnh hiện tại.

Mỗi ảnh được mã hóa được kết hợp với biến số đếm thứ tự ảnh, được ký hiệu là *PicOrderCntVal*.

Khi ảnh hiện tại không phải là ảnh CLVSS, các biến *prevPicOrderCntLsb* và *prevPicOrderCntMsb* được thu nhận như sau:

- Đặt *prevTid0Pic* là ảnh trước đó trong thứ tự giải mã mà có *nuh_layer_id* bằng *nuh_layer_id* của ảnh hiện tại và *TemporalId* bằng 0 và không phải là ảnh RASL hoặc RADL.

- Biến *prevPicOrderCntLsb* được thiết lập bằng **pic_order_cnt_lsb** của *prevTid0Pic*.

- Biến *prevPicOrderCntMsb* được thiết lập bằng *PicOrderCntMsb* của *prevTid0Pic*.

Biến *PicOrderCntMsb* của ảnh hiện tại được thu nhận như sau:

- Nếu ảnh hiện tại là ảnh CLVSS, *PicOrderCntMsb* được thiết lập bằng 0.

- Nếu không phải, *PicOrderCntMsb* được thu nhận như sau:

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

```
PicOrderCntMsb = prevPicOrderCntMsb + MaxPicOrderCntLsb
```

(8-1)

```
else if( ( pic_order_cnt_lsb > prevPicOrderCntLsb ) &&
```

```
(( pic_order_cnt_lsb - prevPicOrderCntLsb ) > (
```

$\text{MaxPicOrderCntLsb} / 2)))$

$\text{PicOrderCntMsb} = \text{prevPicOrderCntMsb} - \text{MaxPicOrderCntLsb}$

else

$\text{PicOrderCntMsb} = \text{prevPicOrderCntMsb}$

PicOrderCntVal được thu nhận như sau:

$$\text{PicOrderCntVal} = \text{PicOrderCntMsb} + \text{pic_order_cnt_lsb} \quad (8-2)$$

Lưu ý 1 – tất cả các ảnh CLVSS sẽ có PicOrderCntVal bằng pic_order_cnt_lsb do đối với các ảnh CLVSS, PicOrderCntMsb được thiết lập bằng 0.

Giá trị của PicOrderCntVal sẽ nằm trong đoạn từ -2^{31} đến $2^{31} - 1$.

Trong một CVS, các giá trị PicOrderCntVal đối với hai ảnh được mã hóa bất kỳ với cùng giá trị của nuh_layer_id sẽ không giống nhau.

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

Hàm $\text{PicOrderCnt}(\text{picX})$ được chỉ rõ như sau:

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

Hàm $\text{DiffPicOrderCnt}(\text{picA}, \text{picB})$ được chỉ rõ như sau:

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

Dòng bit sẽ không chứa dữ liệu mà kết quả trong các giá trị của $\text{DiffPicOrderCnt}(\text{picA}, \text{picB})$ được sử dụng trong xử lý giải mã mà không nằm trong đoạn từ -2^{15} đến $2^{15} - 1$.

Lưu ý 2 - Đặt X là ảnh hiện tại và Y và Z là hai ảnh khác trong cùng CVS, Y và Z được xem là nằm trong cùng hướng thứ tự đầu ra từ X khi cả $\text{DiffPicOrderCnt}(X, Y)$ và $\text{DiffPicOrderCnt}(X, Z)$ đều có giá trị dương hoặc cả hai đều có giá trị âm.

Xử lý giải mã chung để tạo ra các ảnh tham chiếu không khả dụng.

Xử lý này được gọi ra một lần cho mỗi ảnh được mã hóa khi ảnh hiện tại là ảnh CRA với $\text{NoIncorrectPicOutputFlag}$ bằng 1 hoặc ảnh GDR với $\text{NoIncorrectPicOutputFlag}$ bằng 1.

Khi xử lý này được gọi ra, những điều sau sẽ áp dụng:

- Đối với mỗi $\text{RefPicList}[i][j]$, với i trong đoạn từ 0 đến 1, và j trong đoạn từ 0 đến $\text{num_ref_entries}[i][\text{RplsIdx}[i]] - 1$, mà bằng “không có ảnh tham chiếu” ảnh được tạo ra như được quy định trong mục 8.3.4.2 và phần sau đây được áp dụng.

- giá trị của nuh_layer_id đối với ảnh được tạo ra được thiết lập bằng nuh_layer_id của ảnh hiện tại.

- Nếu $\text{st_ref_pic_flag}[i][\text{RplsIdx}[i]][j]$ bằng 1, giá trị của PicOrderCntVal đối với ảnh được tạo ra được thiết lập bằng $\text{RefPicPocList}[i][j]$ và ảnh được tạo ra được đánh dấu là “được sử dụng cho tham chiếu ngắn hạn.”

- Nếu không phải ($\text{st_ref_pic_flag}[i][\text{RplsIdx}[i]][j]$ bằng 0), giá trị của PicOrderCntVal đối với ảnh được tạo ra được thiết lập bằng $\text{RefPicLtPocList}[i][j]$, giá trị của **pic_order_cnt_lsb** đối với ảnh được tạo ra được xem là bằng $(\text{RefPicLtPocList}[i][j] \& (\text{MaxPicOrderCntLsb} - 1))$, và ảnh được tạo ra được đánh dấu là “được sử dụng cho tham chiếu dài hạn.”

- Giá trị của PicOutputFlag đối với ảnh tham chiếu được tạo ra được thiết lập bằng 0.

- $\text{RefPicList}[i][j]$ được thiết lập là ảnh tham chiếu được tạo ra.

Xử lý thu nhận đối với dự đoán vectơ chuyển động độ chói theo thời gian.

Các biến mvLXCol và $\text{availableFlagLXCol}$ được thu nhận như sau:

- Nếu **pic_temporal_mvp_enabled_flag** bằng 0 hoặc $(\text{cbWidth} * \text{cbHeight})$ nhỏ hơn hoặc bằng 32, cả hai thành phần mvLXCol được thiết lập bằng 0 và $\text{availableFlagLXCol}$ được thiết lập bằng 0.

- Nếu không phải (**pic_temporal_mvp_enabled_flag** bằng 1), các bước có thứ tự sau đây được áp dụng.

1. Vectơ chuyển động được sắp xếp cùng vị trí dưới cùng bên phải và các vị trí mẫu biên dưới cùng bên phải được thu nhận như sau:

$$\text{xColBr} = \text{xCb} + \text{cbWidth} \quad (8-421)$$

$$\text{yColBr} = \text{yCb} + \text{cbHeight} \quad (8-422)$$

rightBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?
 SubPicRightBoundaryPos : pic_width_in_luma_samples - 1 (8-423)

botBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?
 SubPicBotBoundaryPos : pic_height_in_luma_samples - 1 (8-424)

- Nếu $yCb \gg CtbLog2SizeY$ bằng $yColBr \gg CtbLog2SizeY$, $yColBr$ nhỏ hơn hoặc bằng $botBoundaryPos$ và $xColBr$ nhỏ hơn hoặc bằng $rightBoundaryPos$, phần sau đây được áp dụng.

- Biến $colCb$ chỉ rõ khối mã hóa độ chói mà bao phủ vị trí được cải biến được đưa ra bởi $((xColBr \gg 3) \ll 3, (yColBr \gg 3) \ll 3)$ bên trong ảnh được sắp xếp cùng vị trí được chỉ rõ bởi $ColPic$.

- Vị trí độ chói $(xColCb, yColCb)$ được thiết lập bằng mẫu trên cùng-bên trái của khối mã hóa độ chói được sắp xếp cùng vị trí được chỉ rõ bởi $colCb$ so với mẫu độ chói trên cùng-bên trái của ảnh được sắp xếp cùng vị trí được chỉ rõ bởi $ColPic$.

- Xử lý thu nhận đối với các vectơ chuyển động được sắp xếp cùng vị trí như được quy định trong điều 8.5.2.12 được gọi ra với $currCb$, $colCb$, $(xColCb, yColCb)$, $refIdxLX$ và $sbFlag$ được thiết lập bằng 0 như là các đầu vào, và đầu ra được gán tới $mvLXCol$ và $availableFlagLXCol$.

- Nếu không phải, cả hai thành phần của $mvLXCol$ được thiết lập bằng 0 và $availableFlagLXCol$ được thiết lập bằng 0.

2. Khi $availableFlagLXCol$ bằng 0, vectơ chuyển động được sắp xếp cùng vị trí trung tâm được thu nhận như sau:

$$xColCtr = xCb + (cbWidth \gg 1) \quad (8-425)$$

$$yColCtr = yCb + (cbHeight \gg 1) \quad (8-426)$$

- Biến $colCb$ chỉ rõ khối mã hóa độ chói mà bao phủ vị trí được cải biến được đưa ra bởi $((xColCtr \gg 3) \ll 3, (yColCtr \gg 3) \ll 3)$ bên trong ảnh được sắp xếp cùng vị trí được chỉ rõ bởi $ColPic$.

- Vị trí độ chói $(xColCb, yColCb)$ được thiết lập bằng mẫu trên cùng-bên trái của khối mã hóa độ chói được sắp xếp cùng vị trí được chỉ rõ bởi $colCb$ so với mẫu độ chói trên cùng-bên trái của ảnh được sắp xếp cùng vị trí được chỉ rõ bởi $ColPic$.

- Xử lý thu nhận đối với các vector chuyển động được sắp xếp cùng vị trí như được quy định trong điều 8.5.2.12 được gọi ra với currCb, colCb, (xColCb, yColCb), refIdxLX và sbFlag được thiết lập bằng 0 như là các đầu vào, và đầu ra được gán tới mvLXCol và availableFlagLXCol.

Xử lý thu nhận đối với các ứng viên hợp nhất theo thời gian dựa trên khối con.

Cờ khả dụng availableFlagSbCol được thu nhận như sau.

- Nếu một hoặc nhiều trong số các điều kiện sau đây là đúng, availableFlagSbCol được thiết lập bằng 0.

- **pic_temporal_mvp_enabled_flag** bằng 0.
- sps_sbtmvp_enabled_flag bằng 0.
- cbWidth nhỏ hơn 8.
- cbHeight nhỏ hơn 8.
- Nếu không phải, các bước có thứ tự sau đây được áp dụng.

1. Vị trí (xCtb, yCtb) của mẫu trên cùng-bên trái của khối cây mã hóa độ chói mà chứa khối mã hóa hiện tại và vị trí (xCtr, yCtr) của mẫu trung tâm phía dưới-bên phải của khối mã hóa độ chói hiện tại được thu nhận như sau:

$$xCtb = (xCb \gg CtuLog2Size) \ll CtuLog2Size \quad (8-542)$$

$$yCtb = (yCb \gg CtuLog2Size) \ll CtuLog2Size \quad (8-543)$$

$$xCtr = xCb + (cbWidth / 2) \quad (8-544)$$

$$yCtr = yCb + (cbHeight / 2) \quad (8-545)$$

2. Vị trí độ chói (xColCtrCb, yColCtrCb) được thiết lập bằng mẫu trên cùng-bên trái của khối mã hóa độ chói được sắp xếp cùng vị trí mà bao phủ vị trí được đưa ra bởi (xCtr, yCtr) bên trong ColPic so với mẫu độ chói trên cùng-bên trái của ảnh được sắp xếp cùng vị trí được chỉ rõ bởi ColPic.

3. Xử lý thu nhận đối với dữ liệu chuyển động cơ sở hợp nhất theo thời gian dựa trên khối con như được quy định trong Điều 8.5.5.4 được gọi ra với vị trí (xCtb, yCtb), vị trí (xColCtrCb, yColCtrCb), cờ khả dụng availableFlagA₁, và cờ sử dụng danh sách dự đoán predFlagLXA₁, và chỉ số tham chiếu refIdxLXA₁, và vector chuyển động mvLXA₁, với X bằng 0 và 1 như là các đầu vào và các vector

chuyển động ctrMvLX , và các cờ sử dụng danh sách dự đoán ctrPredFlagLX của khối được sắp xếp cùng vị trí, với X bằng 0 và 1, và vector chuyển động theo thời gian tempMv như là các đầu ra.

4. Biến $\text{availableFlagSbCol}$ được thu nhận như sau.

- Nếu cả ctrPredFlagL0 và ctrPredFlagL1 bằng 0, $\text{availableFlagSbCol}$ được thiết lập bằng 0.

- Nếu không phải, $\text{availableFlagSbCol}$ được thiết lập bằng 1.

Xử lý thu nhận đối với các ứng viên hợp nhất vector chuyển động điểm điều khiển affin .

Vector chuyển động điểm điều khiển (dưới cùng-bên trái được sắp xếp cùng vị trí) thứ tư $\text{cpMvLXCorner}[3]$, chỉ số tham chiếu $\text{refIdxLXCorner}[3]$, cờ sử dụng danh sách dự đoán $\text{predFlagLXCorner}[3]$ và cờ khả dụng $\text{availableFlagCorner}[3]$ với X 0 và 1 được thu nhận như sau.

- Các chỉ số tham chiếu đối với ứng viên hợp nhất theo thời gian, $\text{refIdxLXCorner}[3]$, với X bằng 0 hoặc 1, được thiết lập bằng 0.

- Các biến mvLXCol và $\text{availableFlagLXCol}$, với X bằng 0 hoặc 1, được thu nhận như sau:

- Nếu **$\text{pic_temporal_mvp_enabled_flag}$** bằng 0, cả hai thành phần của mvLXCol được thiết lập bằng 0 và $\text{availableFlagLXCol}$ được thiết lập bằng 0.

Nếu không phải (**$\text{pic_temporal_mvp_enabled_flag}$** bằng 1), phần sau đây được áp dụng:

$$\text{xColBr} = \text{xCb} + \text{cbWidth} \quad (8-601)$$

$$\text{yColBr} = \text{yCb} + \text{cbHeight} \quad (8-602)$$

- Nếu $\text{yCb} \gg \text{CtbLog2SizeY}$ bằng $\text{yColBr} \gg \text{CtbLog2SizeY}$, yColBr nhỏ hơn $\text{pic_height_in_luma_samples}$ và xColBr nhỏ hơn $\text{pic_width_in_luma_samples}$, phần sau đây được áp dụng.

- Biến colCb chỉ rõ khối mã hóa độ chói mà bao phủ vị trí được cải biến được đưa ra bởi $((\text{xColBr} \gg 3) \ll 3, (\text{yColBr} \gg 3) \ll 3)$ bên trong ảnh được sắp xếp cùng vị trí được chỉ rõ bởi ColPic .

- Vị trí độ chói (xColCb, yColCb) được thiết lập bằng mẫu trên cùng-bên trái của khối mã hóa độ chói được sắp xếp cùng vị trí được chỉ rõ bởi colCb so với mẫu độ chói trên cùng-bên trái của ảnh được sắp xếp cùng vị trí được chỉ rõ bởi ColPic.

- Xử lý thu nhận đối với các vector chuyển động được sắp xếp cùng vị trí như được quy định trong điều 8.5.2.12 được gọi ra với currCb, colCb, (xColCb, yColCb), refIdxLXCorner[3] và sbFlag được thiết lập bằng 0 như là các đầu vào, và đầu ra được gán tới mvLXCol và availableFlagLXCol.

- Nếu không phải, cả hai thành phần của mvLXCol được thiết lập bằng 0 và availableFlagLXCol được thiết lập bằng 0.

- Các biến availableFlagCorner[3], predFlagL0Corner[3], cpMvL0Corner[3] và predFlagL1Corner[3] được thu nhận như sau.

$$\text{availableFlagCorner}[3] = \text{availableFlagL0Col} \quad (8-603)$$

$$\text{predFlagL0Corner}[3] = \text{availableFlagL0Col} \quad (8-604)$$

$$\text{cpMvL0Corner}[3] = \text{mvL0Col} \quad (8-605)$$

$$\text{predFlagL1Corner}[3] = 0 \quad (8-606)$$

Khi slice_type bằng B, các biến availableFlagCorner[3], predFlagL1Corner[3] và cpMvL1Corner[3] được thu nhận như sau.

$$\text{availableFlagCorner}[3] = \text{availableFlagL0Col} \mid \mid \text{availableFlagL1Col} \quad (8-607)$$

$$\text{predFlagL1Corner}[3] = \text{availableFlagL1Col} \quad (8-608)$$

$$\text{cpMvL1Corner}[3] = \text{mvL1Col} \quad (8-609)$$

FIG.11 là sơ đồ giản lược của thiết bị mã video 1100 (ví dụ, bộ mã hóa video 300 hoặc bộ giải mã video 400) theo phương án của sáng chế. Thiết bị mã video 1100 phù hợp để thực hiện các phương án được mô tả ở đây. Thiết bị mã video 1100 bao gồm các cổng vào 1110 và các bộ thu (Rx) 1120 để thu dữ liệu; bộ xử lý, bộ logic, hoặc bộ xử lý trung tâm (CPU-central processing unit) 1130 để xử lý dữ liệu; các bộ truyền (Tx) 1140 và các cổng ra 1150 để truyền dữ liệu; và bộ nhớ 1160 để lưu trữ dữ liệu. Thiết bị mã video 1100 có thể cũng bao gồm các thành phần quang-điện (OE) và các thành phần điện-quang (EO) được ghép nối với các

cổng vào 1110, các đơn vị bộ thu 1120, các đơn vị bộ truyền 1140, và các cổng ra 1150 để đưa ra hoặc đưa vào các tín hiệu quang hoặc các tín hiệu điện.

Bộ xử lý 1130 được thực hiện bởi phần cứng và phần mềm. bộ xử lý 1130 có thể được thực hiện như là một hoặc nhiều chip CPU, các lõi (ví dụ, bộ xử lý đa lõi), mảng cổng khả trình dạng trường (FPGA-field-programmable gate array), mạch tích hợp ứng dụng riêng (ASIC-application specific integrated circuit), và bộ xử lý tín hiệu số (DSP-digital signal processor). Bộ xử lý 1130 truyền thông với các cổng vào 1110, các đơn vị bộ thu 1120, các đơn vị bộ truyền 1140, các cổng ra 1150, và bộ nhớ 1160. Bộ xử lý 1130 bao gồm môđun mã hóa 1170. Môđun mã hóa 1170 thực hiện các phương án được bộc lộ như được mô tả nêu trên. Ví dụ, môđun mã hóa 1170 thực hiện, xử lý, chuẩn bị hoặc cung cấp các chức năng mã hóa-giải mã khác nhau. Sự bao gồm của môđun mã hóa 1170 do đó đưa ra cải thiện căn bản cho chức năng của thiết bị mã video 1100 và tác động tới sự biến đổi của thiết bị mã video 1100 đến trạng thái khác. Ngoài ra, môđun mã hóa 1170 được thực hiện như các lệnh được lưu trữ trong bộ nhớ 1160 và được thực thi bởi bộ xử lý 1130.

Thiết bị mã video 1100 có thể cũng bao gồm các thiết bị đầu vào và/hoặc đầu ra (I/O) 1180 để truyền thông dữ liệu tới và từ người dùng. Các thiết bị I/O 1180 có thể bao gồm các thiết bị đầu ra như màn hình để hiển thị dữ liệu video, các loa để xuất ra dữ liệu audio, v.v. Các thiết bị I/O 1180 có thể cũng 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 này.

Bộ nhớ 1160 bao gồm một hoặc nhiều đĩa, ổ băng và ổ bán dẫn và có thể được sử dụng như một thiết bị lưu trữ dữ liệu tràn dòng, để lưu trữ các chương trình khi các chương trình này được lựa chọn để thực thi và để lưu trữ các lệnh và dữ liệu được đọc trong quá trình thực thi chương trình. Bộ nhớ 1160 có thể là bộ nhớ bất biến và/hoặc không bất biến và có thể là bộ nhớ chỉ đọc (ROM-read-only memory), bộ nhớ truy cập ngẫu nhiên (RAM-random access memory), bộ nhớ định địa chỉ nội dung bậc ba (TCAM-ternary content-addressable memory) và/hoặc bộ nhớ truy cập ngẫu nhiên tĩnh (SRAM-static random-access memory).

FIG.12 là sơ đồ giản lược của phương án về phương tiện mã 1200. Trong phương án của sáng chế, phương tiện mã 1200 được thực hiện trong thiết bị mã video 1202 (ví dụ, bộ mã hóa video 300 hoặc bộ giải mã video 400). Thiết bị mã video 1202 bao gồm phương tiện thu 1201. Phương tiện thu 1201 có cấu trúc để

thu ảnh để mã hóa hoặc thu dòng bit để giải mã. Thiết bị mã video 1202 bao gồm phương tiện truyền 1207 được ghép nối với phương tiện thu 1201. Phương tiện truyền 1207 có cấu trúc để truyền dòng bit tới bộ giải mã hoặc truyền ảnh được giải mã tới phương tiện hiển thị (ví dụ, một trong số các thiết bị I/O 1180).

Thiết bị mã video 1202 bao gồm phương tiện lưu trữ 1203. Phương tiện lưu trữ 1203 được ghép nối tới ít nhất một trong số phương tiện thu 1201 hoặc phương tiện truyền 1207. Phương tiện lưu trữ 1203 có cấu trúc để lưu trữ các lệnh. Thiết bị mã video 1202 cũng bao gồm phương tiện xử lý 1205. Phương tiện xử lý 1205 được ghép nối tới phương tiện lưu trữ 1203. Phương tiện xử lý 1205 có cấu trúc để thực thi các lệnh được lưu trữ trong phương tiện lưu trữ 1203 để thực hiện các phương pháp được bộc lộ ở đây.

Cũng cần được hiểu rằng các bước của các phương pháp ví dụ được mô tả ở đây không cần thiết được yêu cầu được thực hiện trong thứ tự được mô tả, và thứ tự của các bước của các phương pháp này sẽ được hiểu chỉ là ví dụ. Tương tự, các bước bổ sung có thể được chứa trong các phương pháp này, và các bước nhất định có thể được bỏ qua hoặc được kết hợp, trong các phương pháp đồng nhất với các phương án khác nhau của sáng chế.

Trong khi một vài phương án đã được mô tả trong sáng chế, sẽ được hiểu rằng các hệ thống và phương pháp được bộc lộ có thể được thực hiện trong nhiều dạng cụ thể khác mà không đi chệch khỏi bản chất hoặc phạm vi của sáng chế. Các ví dụ được xem xét như là minh họa và không bị giới hạn, và mục đích không bị giới hạn ở các chi tiết được mô tả ở đây. Ví dụ, các bộ phận hoặc các thành phần khác nhau có thể được kết hợp hoặc được tích hợp vào hệ thống khác, hoặc một vài đặc điểm có thể được bỏ qua hoặc không được thực hiện.

Ngoài ra, các kỹ thuật, hệ thống, hệ thống con, và phương pháp được mô tả và được minh họa trong các phương án khác nhau như là thành phần rời rạc hoặc riêng biệt có thể được kết hợp hoặc được tích hợp với các hệ thống, các mô đun, các kỹ thuật, hoặc các phương pháp khác mà không đi chệch khỏi phạm vi của sáng chế. Các hạng mục khác được mô tả hoặc được thể hiện như là được ghép nối hoặc được ghép nối trực tiếp hoặc truyền thông với nhau có thể được ghép nối gián tiếp hoặc truyền thông qua một vài giao diện, thiết bị, hoặc thành phần trung gian bằng điện, cơ học, hoặc cách thức khác. Các ví dụ khác về các thay đổi, thay thế, và bổ sung có thể được thực hiện bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật

và có thể được thực hiện mà không đi chệch khỏi bản chất và phạm vi được bộc lộ ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã được thực hiện bởi bộ giải mã video, bao gồm:

thu dòng bit video mà bao gồm đơn vị lớp trừu tượng mạng (NAL - network abstraction layer) thông tin tiêu đề ảnh (PH - picture header) mà có loại đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL - video coding layer) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID) lớp của đơn vị NAL thông tin tiêu đề ảnh bằng ID lớp của đơn vị NAL lớp mã hóa video được liên kết và ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh bằng ID theo thời gian của đơn vị NAL lớp mã hóa video được liên kết; và

giải mã ảnh được mã hóa từ dòng bit video để thu nhận ảnh được giải mã.

2. Phương pháp theo điểm 1, còn bao gồm nhận dạng lớp của ảnh được mã hóa dựa trên ID lớp của đơn vị NAL thông tin tiêu đề ảnh, và nhận dạng vị trí theo thời gian của ảnh được mã hóa trong dòng bit video dựa trên ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh.

3. Phương pháp theo điểm 1, trong đơn vị NAL lớp mã hóa video được liên kết bao gồm lát của ảnh được mã hóa.

4. Phương pháp theo điểm 1, trong đó đơn vị NAL thông tin tiêu đề ảnh là đơn vị NAL không phải VCL.

5. Phương pháp theo điểm 1, trong đó ID lớp của đơn vị NAL thông tin tiêu đề ảnh được chỉ định là `nuh_layer_id`.

6. Phương pháp theo điểm 1, trong đó ID lớp của đơn vị NAL lớp mã hóa video nhận dạng lớp mà bao gồm đơn vị NAL lớp mã hóa video được liên kết.

7. Phương pháp theo điểm bất kỳ trong số các điểm 1 đến 6, còn bao gồm hiển thị ảnh được giải mã trên màn hình của thiết bị điện tử.

8. Phương pháp mã hóa được thực hiện bởi bộ mã hóa video, phương pháp này bao gồm:

tạo ra đơn vị lớp trừu tượng mạng (NAL - network abstraction layer) thông tin tiêu đề ảnh (PH - picture header) mà có loại đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL - video coding layer) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID) lớp của đơn vị NAL thông tin

tiêu đề ảnh bằng ID lớp của đơn vị NAL lớp mã hóa video được liên kết và ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh bằng ID theo thời gian của đơn vị NAL lớp mã hóa video được liên kết;

mã hóa đơn vị NAL thông tin tiêu đề ảnh và đơn vị NAL lớp mã hóa video được liên kết vào dòng bit video; và

lưu trữ dòng bit video để truyền thông tới bộ giải mã video.

9. Phương pháp theo điểm 8, trong đó ID lớp của đơn vị NAL thông tin tiêu đề ảnh nhận dạng lớp của ảnh được mã hóa trong dòng bit video, và trong đó ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh nhận dạng vị trí theo thời gian của ảnh được mã hóa trong dòng bit đa lớp.

10. Phương pháp theo điểm 8 hoặc 9, trong đó đơn vị NAL lớp mã hóa video được liên kết bao gồm lát của ảnh được mã hóa, và trong đó đơn vị NAL thông tin tiêu đề ảnh là đơn vị NAL không phải VCL.

11. Thiết bị giải mã, bao gồm:

bộ thu có cấu trúc để thu dòng bit video mà bao gồm đơn vị lớp trừu tượng mạng (NAL - network abstraction layer) thông tin tiêu đề ảnh (PH - picture header) mà có loại đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL - video coding layer) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID) lớp của đơn vị NAL thông tin tiêu đề ảnh bằng ID lớp của đơn vị NAL lớp mã hóa video được liên kết và ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh bằng ID theo thời gian của đơn vị NAL lớp mã hóa video được liên kết;

bộ nhớ được ghép nối tới bộ thu, bộ nhớ này lưu trữ các lệnh; và

bộ xử lý được ghép nối tới bộ nhớ, bộ xử lý có cấu trúc để thực thi các lệnh để làm cho thiết bị giải mã:

giải mã ảnh được mã hóa từ dòng bit video để thu nhận ảnh được giải mã.

12. Thiết bị giải mã theo điểm 11, trong đó bộ xử lý có cấu trúc để thực thi các lệnh để làm cho thiết bị giải mã nhận dạng lớp của ảnh được mã hóa trong dòng bit video dựa trên ID lớp của đơn vị NAL thông tin tiêu đề ảnh, và nhận dạng vị trí theo thời

gian của ảnh được mã hóa trong dòng bit video dựa trên ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh.

13. Thiết bị giải mã theo điểm 11 hoặc 12, trong đó đơn vị NAL lớp mã hóa video được liên kết bao gồm lát của ảnh được mã hóa.

14. Thiết bị mã hóa, bao gồm:

bộ nhớ mà chứa các lệnh;

bộ xử lý được ghép nối tới bộ nhớ, bộ xử lý có cấu trúc để thực hiện các lệnh để làm cho thiết bị mã hóa:

tạo ra đơn vị lớp trừu tượng mạng (NAL - network abstraction layer) thông tin tiêu đề ảnh (PH - picture header) mà có loại đơn vị NAL thông tin tiêu đề ảnh (PH_NUT) và đơn vị NAL lớp mã hóa video (VCL - video coding layer) được liên kết, PH_NUT biểu thị rằng ký hiệu nhận dạng (ID) lớp của đơn vị NAL thông tin tiêu đề ảnh bằng ID lớp của đơn vị NAL lớp mã hóa video được liên kết và ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh bằng ID theo thời gian của đơn vị NAL lớp mã hóa video được liên kết; mã hóa đơn vị NAL thông tin tiêu đề ảnh và đơn vị NAL lớp mã hóa video được liên kết vào dòng bit video; và

bộ truyền được ghép nối tới bộ xử lý, bộ truyền có cấu trúc để truyền dòng bit video tới bộ giải mã video.

15. Thiết bị mã hóa theo điểm 14, trong đó ID lớp của đơn vị NAL thông tin tiêu đề ảnh nhận dạng lớp của ảnh được mã hóa trong dòng bit video, và trong đó ID theo thời gian của đơn vị NAL thông tin tiêu đề ảnh nhận dạng vị trí theo thời gian của ảnh được mã hóa trong dòng bit video.

16. Thiết bị mã, bao gồm:

bộ thu được có cấu trúc để thu ảnh để mã hóa hoặc thu dòng bit để giải mã.

bộ truyền được ghép nối tới bộ thu, bộ truyền có cấu trúc để truyền dòng bit tới bộ giải mã hoặc để truyền ảnh được giải mã tới màn hình;

bộ nhớ được ghép nối tới ít nhất một trong số bộ thu hoặc bộ truyền, bộ nhớ có cấu trúc để lưu trữ các lệnh; và

bộ xử lý được ghép nối tới bộ nhớ, bộ xử lý có cấu trúc để thực thi các lệnh được lưu trữ trong bộ nhớ để thực hiện phương pháp theo điểm bất kỳ trong số các điểm 1 đến 7 hoặc phương pháp theo điểm bất kỳ trong số các điểm 8 đến 10.

17. Thiết bị mã theo điểm 16, còn bao gồm màn hình để hiển thị ảnh được giải mã.

18. Hệ thống, bao gồm:

bộ mã hóa, và:

bộ giải mã truyền thông với bộ mã hóa,

trong đó bộ mã hóa bao gồm thiết bị mã hóa theo điểm bất kỳ trong số các điểm 14 hoặc 15, và bộ giải mã bao gồm thiết bị giải mã theo điểm bất kỳ trong số các điểm 11 đến 13.

19. Phương tiện mã, bao gồm:

phương tiện thu có cấu trúc để thu ảnh để mã hóa hoặc thu dòng bit để giải mã.

phương tiện truyền được ghép nối tới phương tiện thu, phương tiện truyền có cấu trúc để truyền dòng bit tới phương tiện giải mã hoặc để truyền ảnh được giải mã tới phương tiện hiển thị;

phương tiện lưu trữ được ghép nối tới ít nhất một trong số phương tiện thu hoặc phương tiện truyền, phương tiện lưu trữ có cấu trúc để lưu trữ các lệnh; và

phương tiện xử lý được ghép nối tới phương tiện lưu trữ, phương tiện xử lý có cấu trúc để thực thi các lệnh được lưu trữ trong phương tiện lưu trữ để thực hiện phương pháp theo điểm bất kỳ trong số các điểm 1 đến 7 hoặc phương pháp theo điểm bất kỳ trong số các điểm 8 đến 10.

20. Phương tiện lưu trữ đọc được bởi máy tính, trong đó phương tiện lưu trữ đọc được bởi máy tính lưu trữ chương trình máy tính mà có thể được thực thi bởi bộ xử lý, và khi chương trình máy tính được thực thi bởi bộ xử lý, bộ xử lý thực hiện phương pháp theo điểm bất kỳ trong số các điểm 1 đến 7 hoặc phương pháp theo điểm bất kỳ trong số các điểm 8 đến 10.

21. Phương tiện lưu trữ đọc được bởi máy tính, trong đó phương tiện lưu trữ đọc được bởi máy tính lưu trữ dòng bit thu được bằng cách sử dụng phương pháp theo điểm bất kỳ trong số các điểm 8 đến 10.

22. Bộ mã, bao gồm hệ mạch xử lý để thực hiện phương pháp theo điểm bất kỳ trong số các điểm 1 đến 7 hoặc phương pháp theo điểm bất kỳ trong số các điểm 8 đến 10.

23. Hệ thống, bao gồm:

bộ mã hóa, và:

bộ giải mã truyền thông với bộ mã hóa, trong đó bộ giải mã hoặc bộ mã hóa bao gồm thiết bị mã theo điểm 16 hoặc 17.

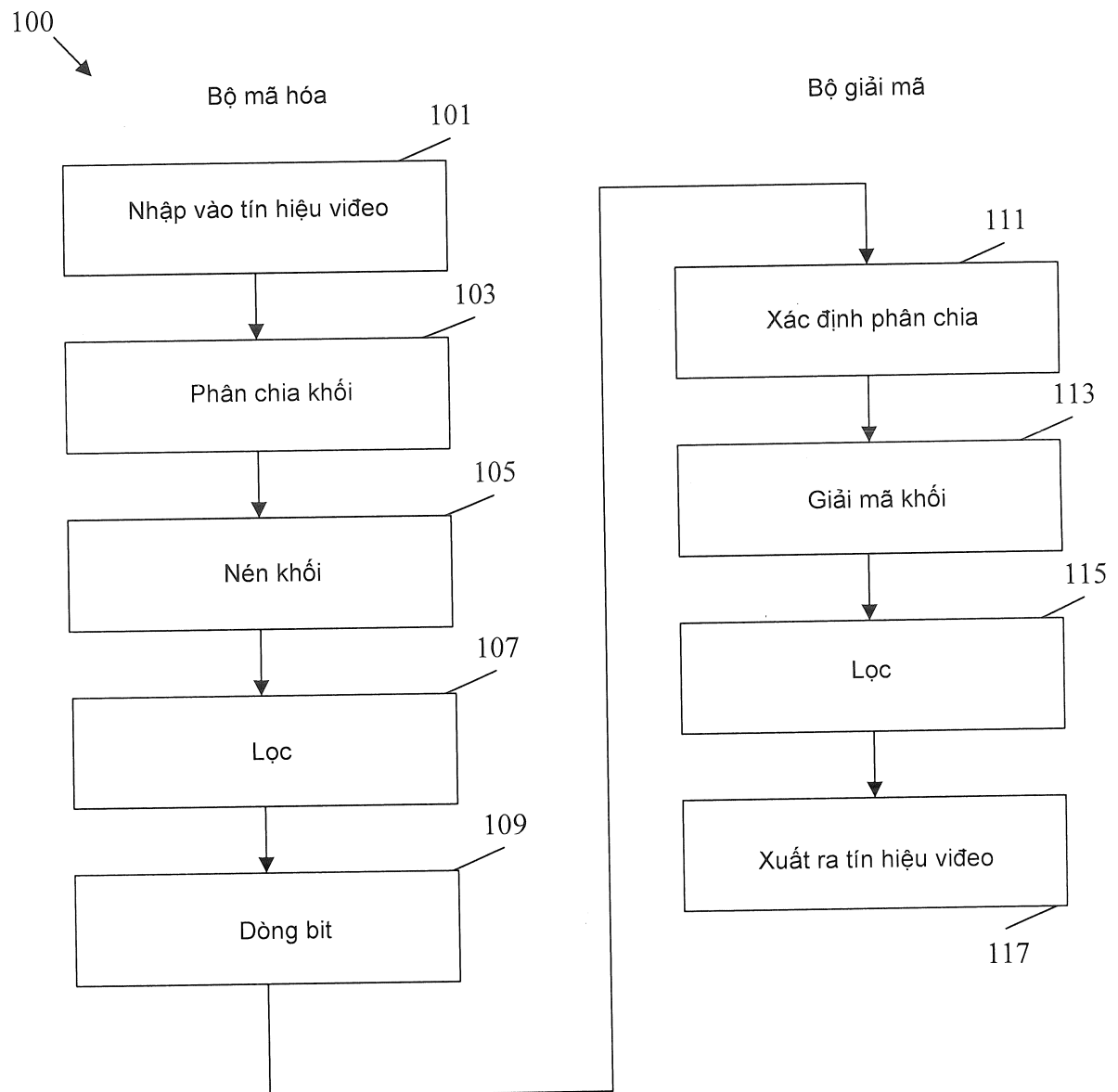


FIG. 1

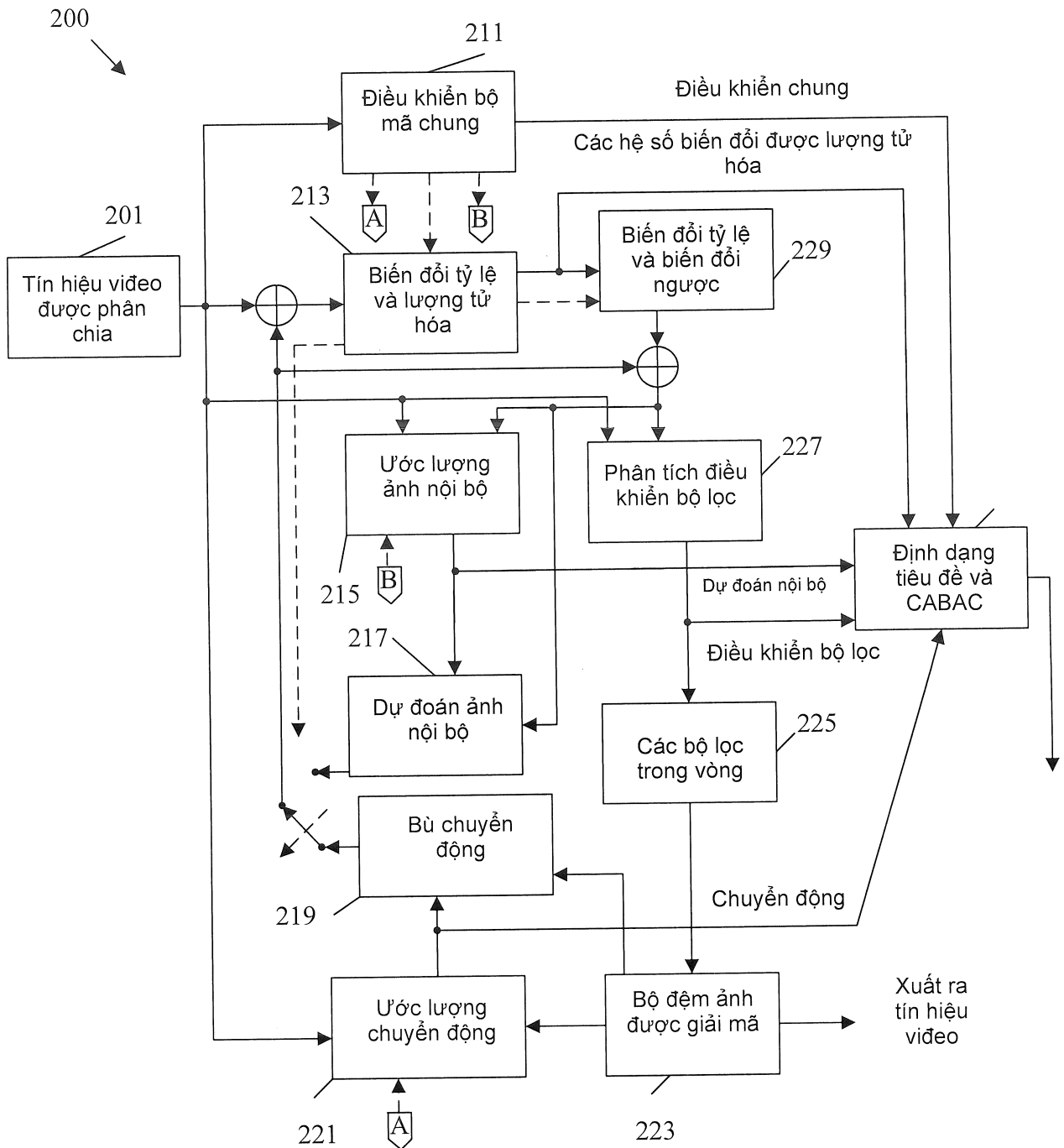


FIG. 2

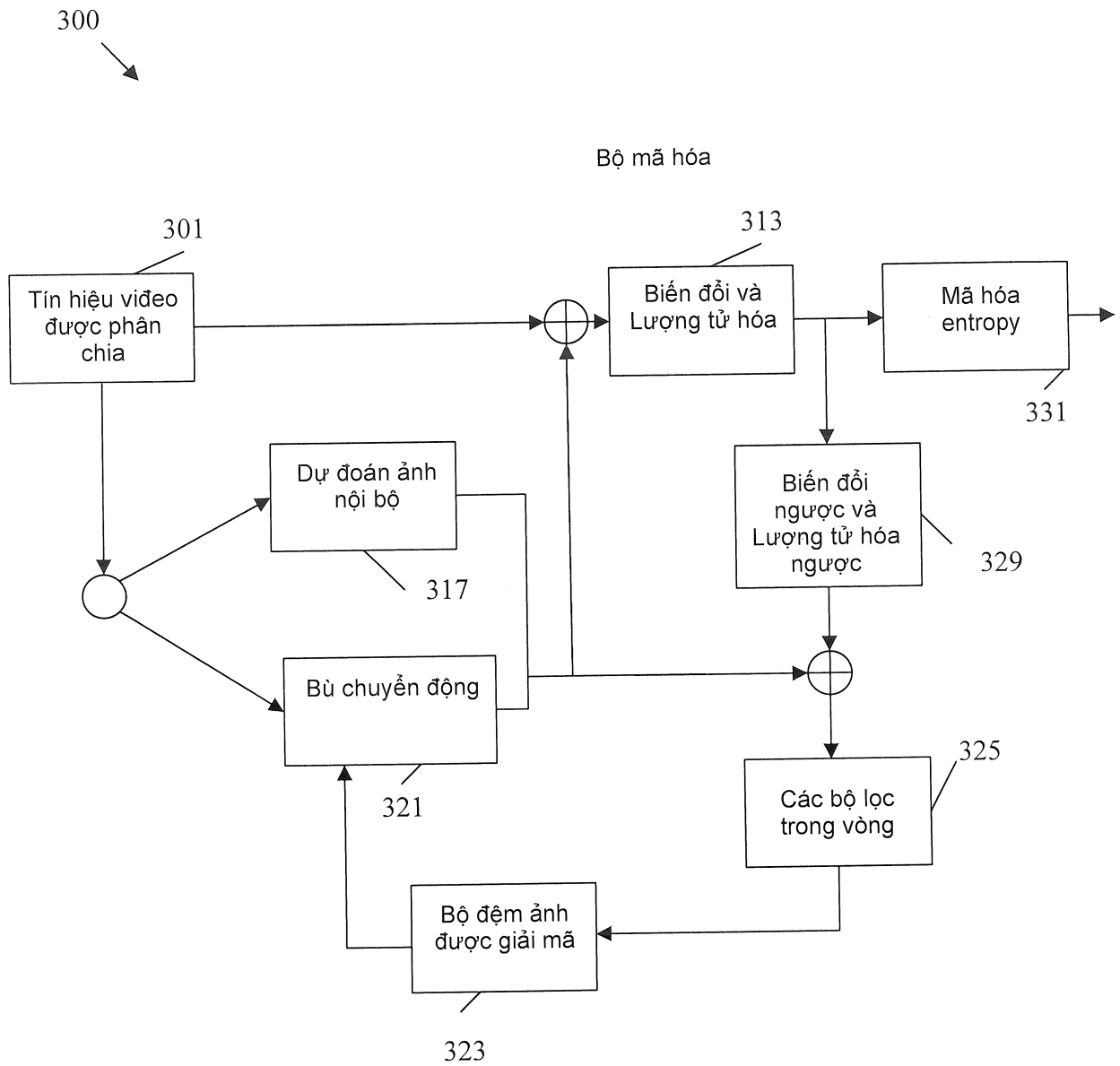


FIG. 3

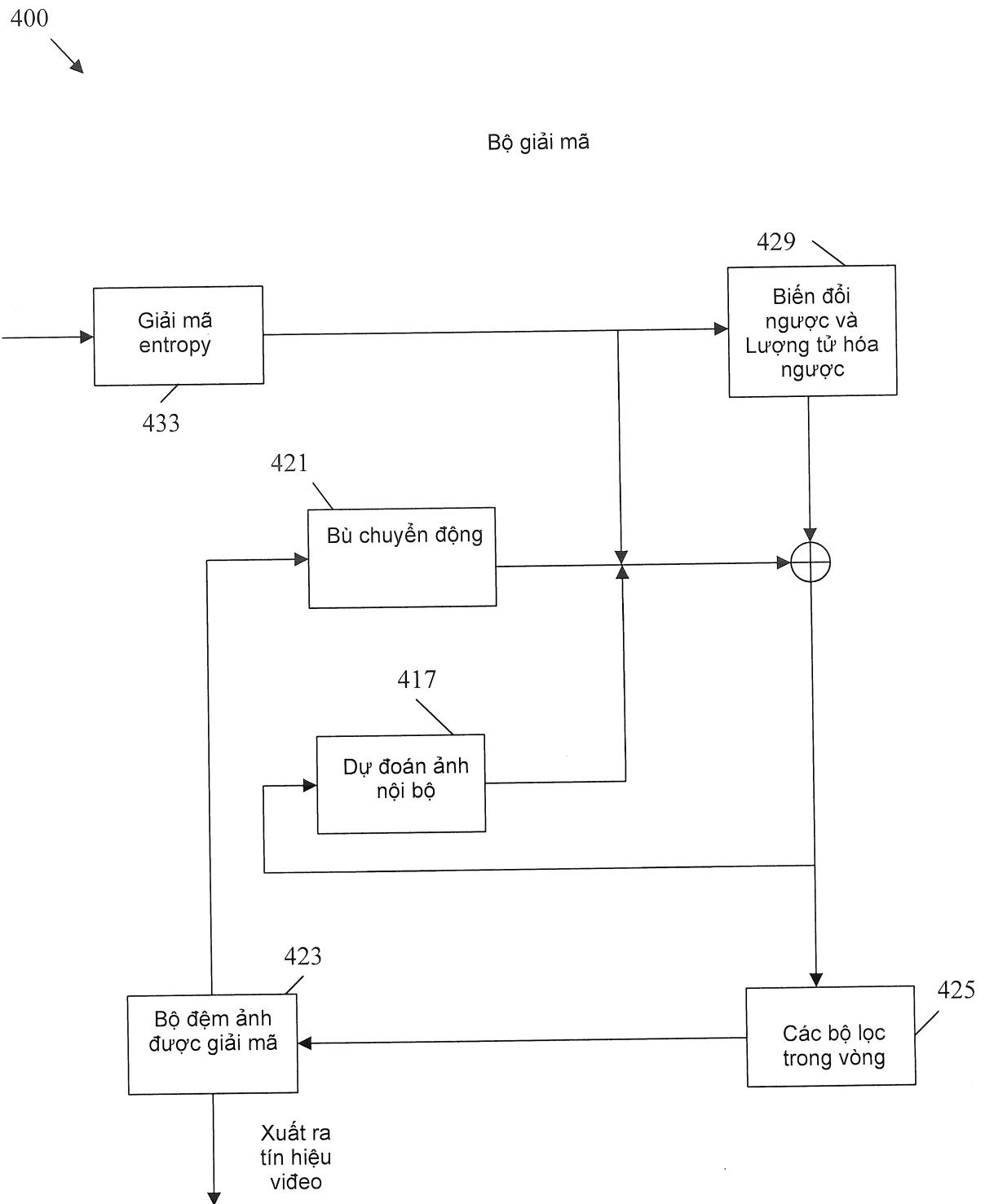


FIG. 4

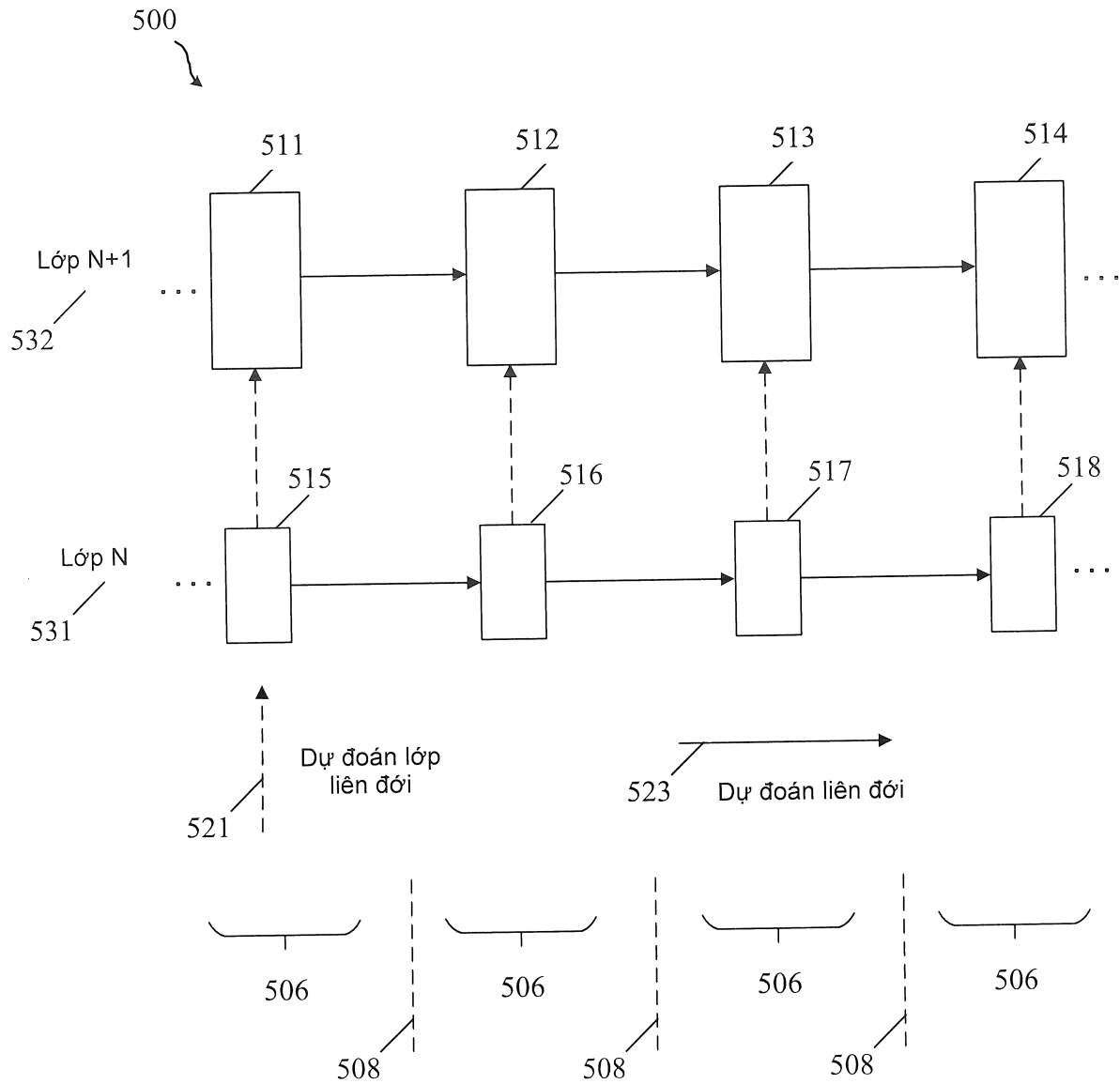


FIG. 5

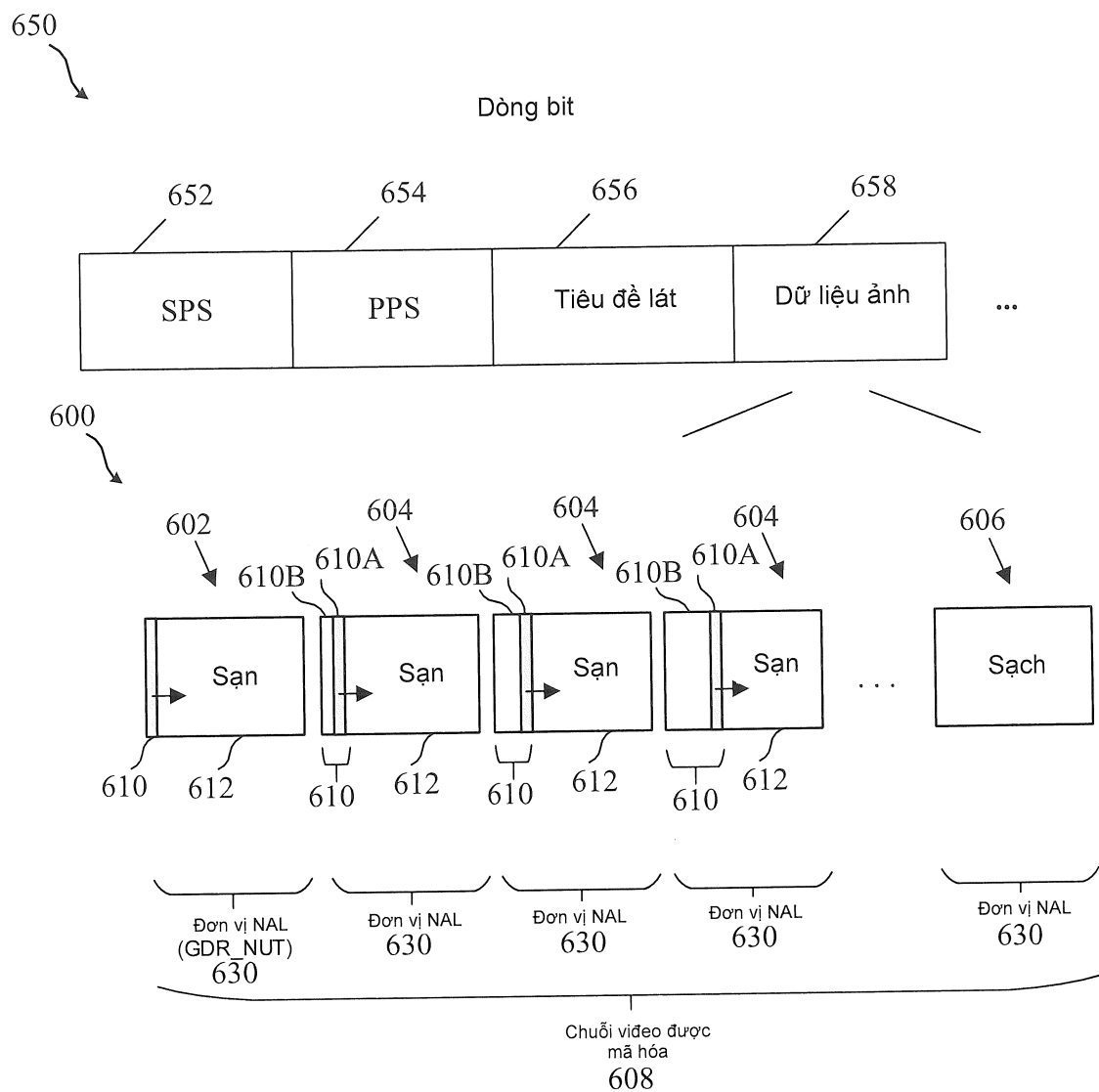


FIG. 6

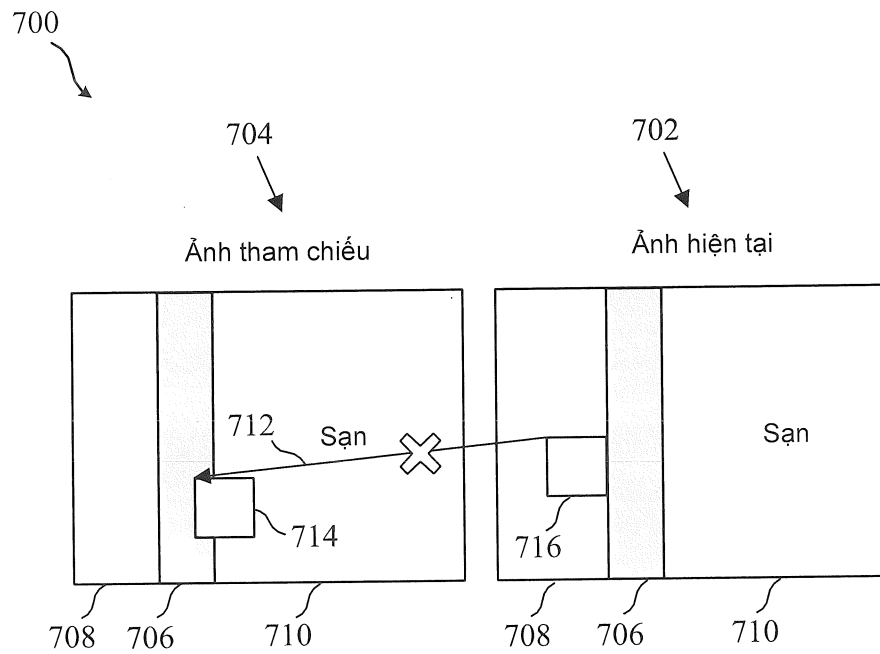


FIG. 7

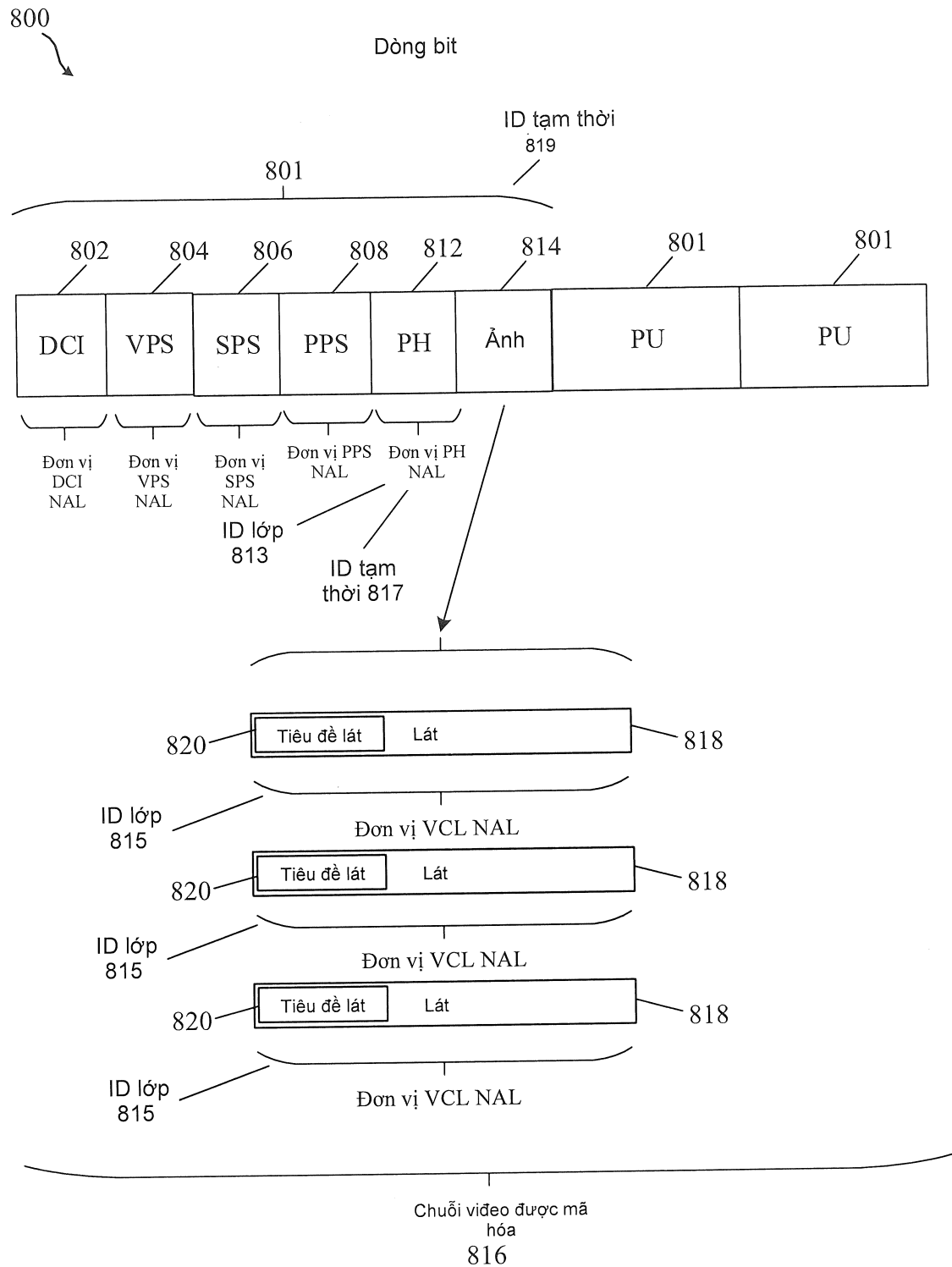


FIG. 8

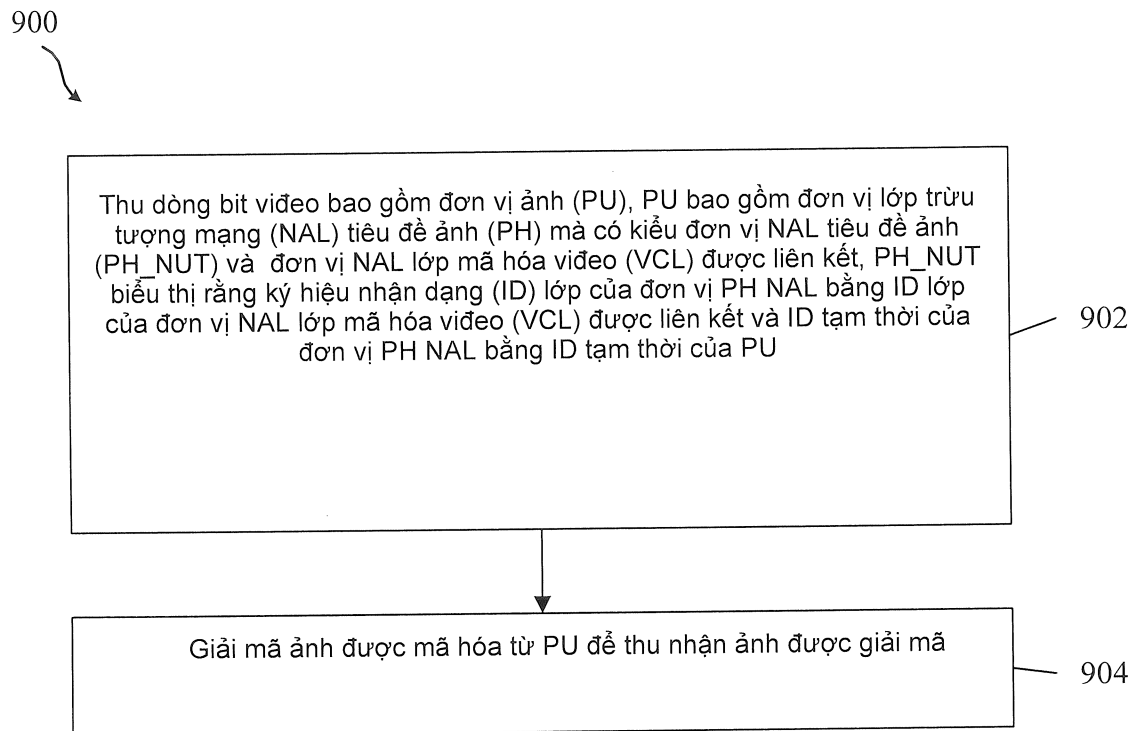


FIG. 9

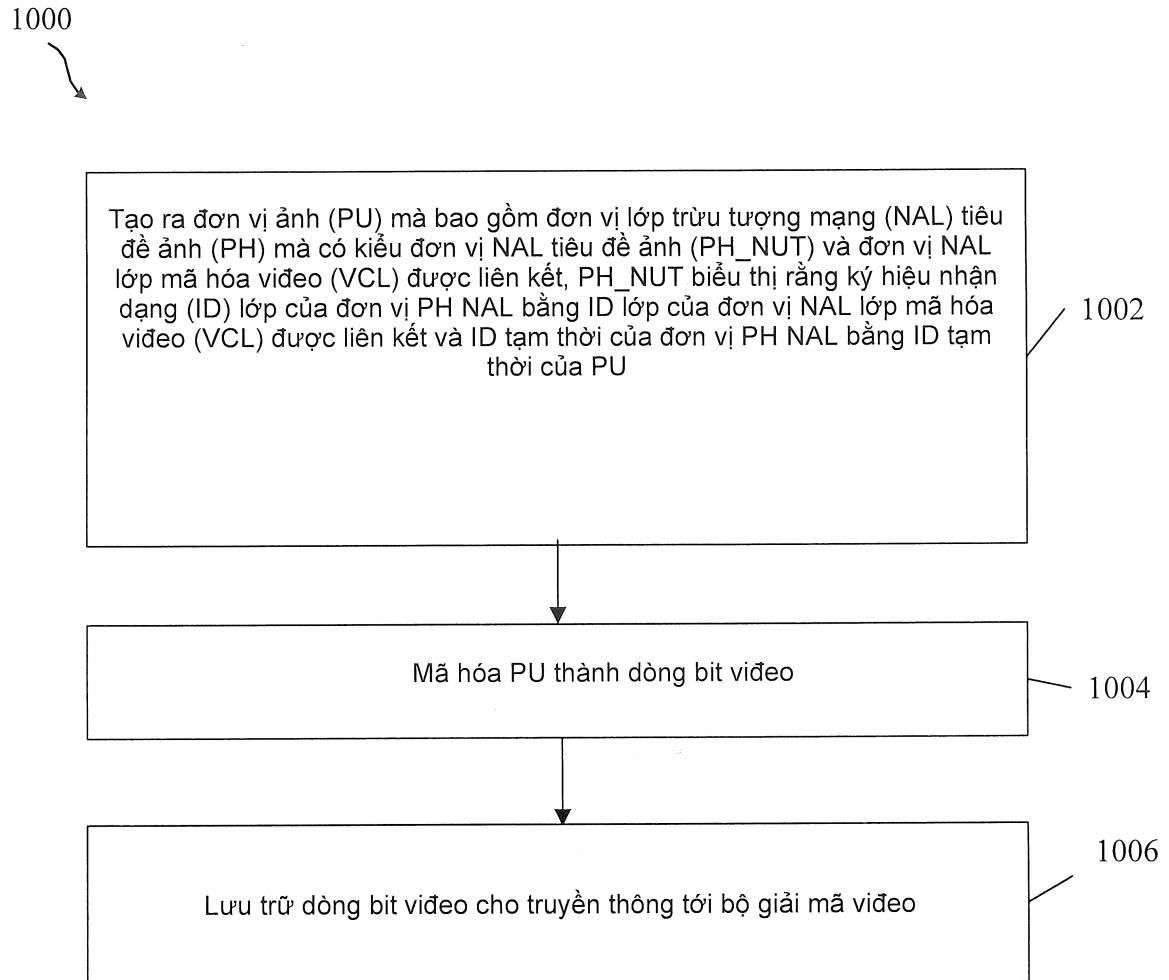


FIG. 10

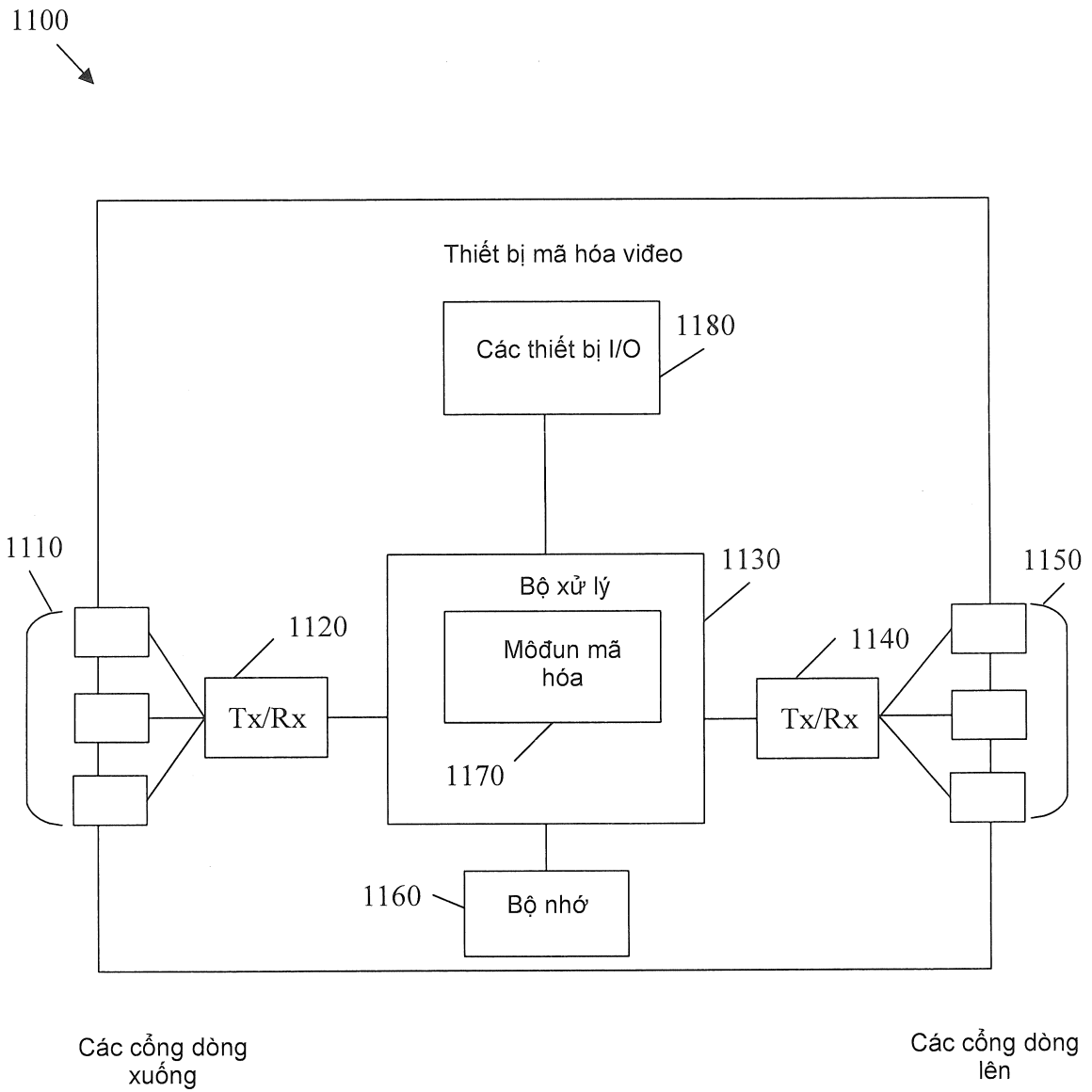


FIG. 11

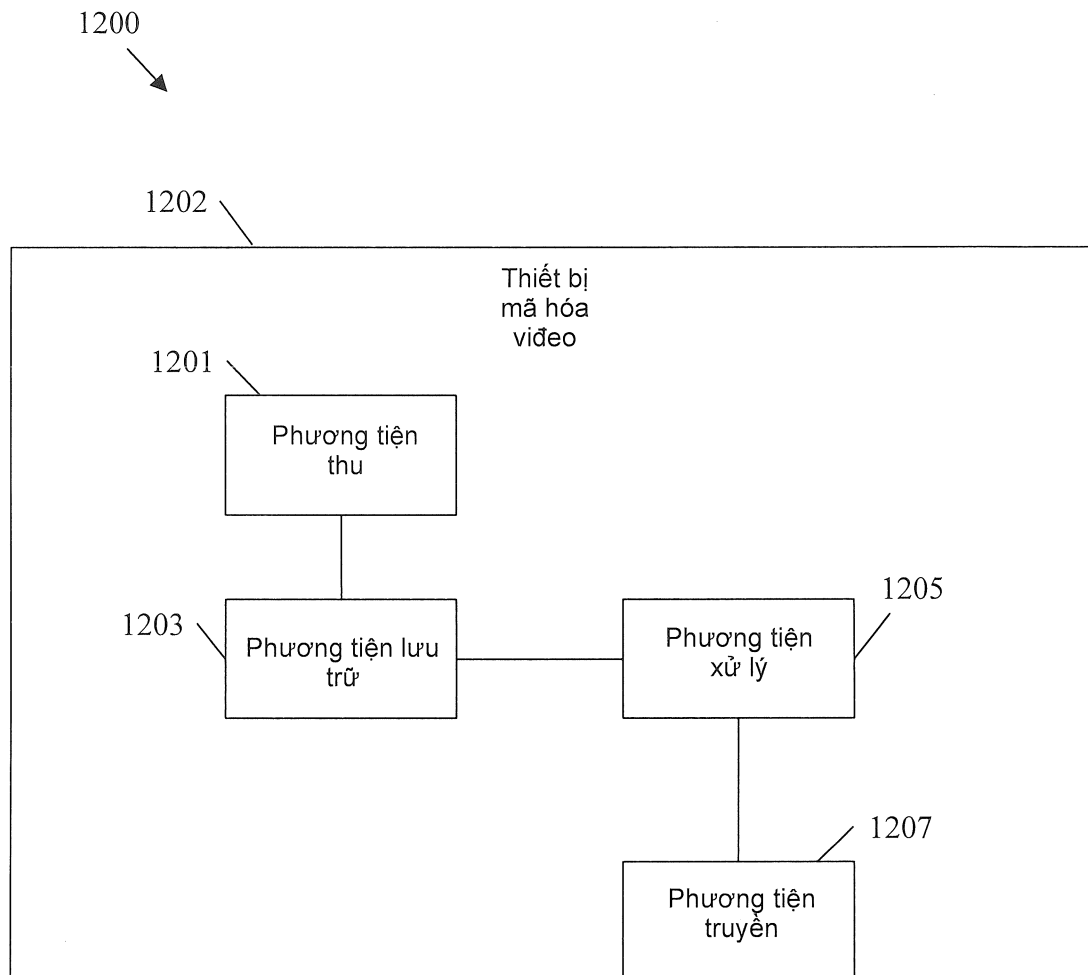


FIG. 12