



- (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
1-0052288
(51)^{2018.01} **G06F 15/16; H04N 19/176; H04N 19/46; H04N 19/167** (13) **B**

-
- (21) 1-2022-02567 (22) 17/09/2020
(86) PCT/US2020/051313 17/09/2020 (87) WO 2021/061493 01/04/2021
(30) 62/905,236 24/09/2019 US
(45) 27/10/2025 451 (43) 25/07/2022 412A
(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)
Huawei Administration Building, Bantian, Longgang District, Shenzhen, Guangdong
518129, China
(72) WANG, Ye-Kui (US).
(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)
-

- (54) PHƯƠNG PHÁP ĐƯỢC THỰC HIỆN BỞI BỘ GIẢI MÃ, PHƯƠNG PHÁP MÃ HÓA ĐƯỢC THỰC HIỆN BỞI BỘ MÃ HÓA, THIẾT BỊ GIẢI MÃ VÀ MÃ HÓA, PHƯƠNG TIỆN LƯU TRỮ ĐƯỢC ĐỌC ĐƯỢC BẰNG MÁY TÍNH, PHƯƠNG TIỆN LƯU TRỮ KHÔNG TẠM THỜI, VÀ PHƯƠNG TIỆN ĐỌC ĐƯỢC BỞI MÁY TÍNH KHÔNG TẠM THỜI

(21) 1-2022-02567

(57) Sáng chế đề cập đến cơ chế lập hóa video, cơ chế này bao gồm phương pháp được thực hiện bởi bộ giải mã, thiết bị giải mã và phương tiện không tạm thời đọc được bởi máy tính. Cơ chế này bao gồm việc mã hóa hình ảnh được lập mã trong một hoặc nhiều đơn vị lớp trừu tượng hóa mạng loại lớp lập mã video (video coding layer (VCL) network abstraction layer (NAL) unit) trong dòng bit. Đơn vị non-VCL NAL được mã hóa thành dòng bit sao cho mã định danh theo thời gian (TemporalId) cho đơn vị non-VCL NAL được ràng buộc bằng TemporalId của đơn vị truy cập (access unit, AU) chứa đơn vị non-VCL NAL khi kiểu đơn vị NAL (nal_unit_type) của non-VCL NAL biểu thị thông điệp thông tin tăng cường bổ sung (supplemental enhancement information, SEI). Tập hợp các phép kiểm tra sự tương thích dòng bit được thực hiện trên dòng bit dựa vào thông điệp SEI. Dòng bit này được lưu trữ để truyền thông về phía bộ giải mã.

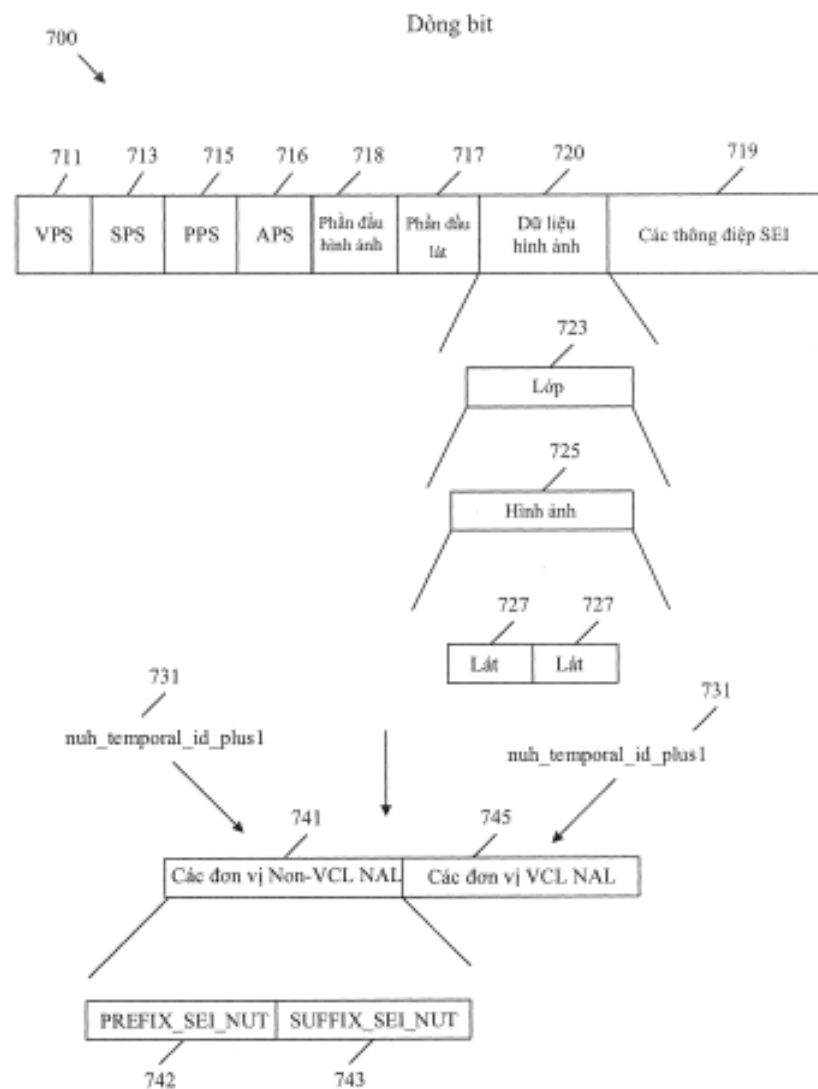


FIG. 7

Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế đề cập đến việc lập mã video, và cụ thể là liên quan đến các cải tiến về các tham số báo hiệu để hỗ trợ việc lập mã các dòng bit đa lớp.

Tình trạng kỹ thuật của sáng chế

Lượng dữ liệu video cần thiết để thể hiện video kể cả tương đối ngắn có thể vẫn là đáng kể, có thể dẫn đến các khó khăn khi dữ liệu được tạo dòng hoặc theo cách khác, được truyền thông qua mạng truyền thông với dung lượng băng thông giới hạn. Do đó, dữ liệu video thường được nén trước khi được truyền thông qua các mạng viễn thông hiện nay. Kích thước của video cũng có thể là một vấn đề khi video được lưu trữ trên thiết bị lưu trữ vì các tài nguyên bộ nhớ có thể bị giới hạn. Các thiết bị nén video thường sử dụng phần mềm và/hoặc phần cứng ở nguồn để lập mã dữ liệu video trước khi truyền dẫn hoặc lưu trữ, nhờ đó làm giảm lượng dữ liệu cần thiết để biểu diễn các hình ảnh video số. Sau đó, dữ liệu nén được nhận ở đích bởi thiết bị giải nén video sẽ giải mã dữ liệu video này. Với các tài nguyên mạng bị giới hạn và nhu cầu chất lượng video cao hơn ngày càng tăng, mong muốn có các kỹ thuật nén và giải nén cải tiến nâng cao tỷ lệ nén mà làm giảm ít hoặc không giảm chất lượng hình ảnh.

Bản chất kỹ thuật của sáng chế

Theo một phương án, sáng chế đề xuất phương pháp được thực hiện bởi bộ giải mã, phương pháp này bao gồm các bước: nhận, bởi bộ thu của bộ giải mã, dòng bit bao gồm hình ảnh được lập mã trong một hoặc nhiều đơn vị lớp trừu tượng hóa mạng loại lớp lập mã video (video coding layer (VCL) network abstraction layer (NAL) unit) và đơn vị non-VCL NAL, trong đó mã định danh theo thời gian (TemporalId) cho đơn vị non-VCL NAL được ràng buộc bằng TemporalId của đơn vị truy cập (access unit, AU) chứa đơn vị non-VCL NAL khi kiểu đơn vị NAL (nal_unit_type) của non-VCL NAL bằng kiểu đơn vị NAL thông tin tăng cường bổ sung (supplemental enhancement

information (SEI) NAL unit type) tiền tố (PREFIX_SEI_NUT) hoặc kiểu đơn vị SEI NAL hậu tố (SUFFIX_SEI_NUT); suy ra, bởi bộ xử lý của bộ giải mã, TemporalId cho đơn vị non-VCL NAL dựa vào phần tử cú pháp mã định danh theo thời gian phần đầu đơn vị NAL cộng một ($\text{nuh_temporal_id_plus1}$) trong đơn vị non-VCL NAL; và giải mã, bởi bộ xử lý của bộ giải mã, hình ảnh được lập mã từ các đơn vị VCL NAL để tạo ra hình ảnh được giải mã.

Một chuỗi video có thể bao gồm nhiều hình ảnh. Để bảo đảm các hình ảnh được hiển thị theo thứ tự đúng, các hệ thống lập mã video có thể gán các hình ảnh này cho TemporalId. Một số hệ thống lập mã video sử dụng các lớp của các hình ảnh, trong đó mỗi lớp về cơ bản bao gồm cùng video của các độ phân giải, các kích thước hình ảnh, các tốc độ khung hình v.v. khác nhau. Các hình ảnh trong các lớp khác nhau có thể được hiển thị theo cách luân phiên, tùy thuộc vào các điều kiện ở bộ giải mã. Do đó, các hình ảnh trong các lớp khác nhau được định vị ở cùng điểm trong chuỗi video chia sẻ cùng TemporalId. Ngoài ra, các hình ảnh trong các lớp khác nhau chia sẻ cùng TemporalId tạo ra AU. Ví dụ, bộ giải mã có thể hiển thị một hình ảnh được chọn từ một lớp ở mỗi AU để hiển thị một chuỗi video. Một số hệ thống lập mã video sử dụng các thông điệp SEI. Thông điệp SEI chứa thông tin không cần thiết cho quy trình giải mã để xác định các giá trị của các mẫu trong các hình ảnh được giải mã. Ví dụ, các thông điệp SEI có thể chứa các tham số được sử dụng bởi bộ giải mã tham chiếu giả định (hypothetical reference decoder, HRD) hoạt động ở bộ mã hóa để kiểm tra dòng bit về sự tương thích với các chuẩn. Ngoài ra, các hệ thống lập mã video có thể lập mã chuỗi video thành dòng bit dưới dạng các lớp của các hình ảnh. Các thông điệp SEI có thể có liên quan đến các hình ảnh khác nhau và/hoặc các tổ hợp khác nhau của các lớp. Do đó, việc bảo đảm rằng thông điệp SEI thích hợp được liên kết với các hình ảnh/các lớp thích hợp trở thành thách thức trong các dòng bit đa lớp phức tạp. Trong trường hợp thông điệp SEI không được liên kết với lớp/hình ảnh đúng, HRD có thể không thể kiểm tra lớp/hình ảnh về sự tương thích theo cách đúng đắn. Việc này có thể dẫn đến các sai số mã hóa.

Ví dụ này bao gồm cơ chế để liên kết đúng các thông điệp SEI với các hình ảnh/các lớp tương ứng. Các dòng bit đa lớp có thể tổ chức các hình ảnh và các tham số có liên quan thành các AU. AU là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian đầu ra. Thông điệp SEI có thể được định

vị trong cùng AU dưới dạng hình ảnh thứ nhất được liên kết với thông điệp SEI. Ngoài ra, thông điệp SEI còn được gán TemporalId. TemporalId là mã định danh biểu thị vị trí tương đối của đơn vị NAL trong chuỗi video. TemporalId của thông điệp SEI được ràng buộc bằng TemporalId của AU chứa thông điệp SEI. Nói cách khác, các hình ảnh có trong các đơn vị VCL NAL và các tham số có trong các đơn vị non-VCL NAL. Khi đơn vị non-VCL NAL là đơn vị SEI NAL chứa thông điệp SEI, TemporalId của đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL này. Phương pháp này bảo đảm rằng các thông điệp SEI được liên kết đúng với các hình ảnh tương ứng trong các AU. Do đó, có thể tránh được các lỗi khác nhau. Kết quả là, chức năng của bộ mã hóa và bộ giải mã được nâng cao. Ngoài ra, hiệu quả lập mã có thể tăng lên, giảm việc sử dụng tài nguyên báo hiệu bộ xử lý, bộ nhớ và/hoặc mạng ở cả bộ mã hóa lẫn bộ giải mã.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó nal_unit_type của non-VCL NAL bằng PREFIX_SEI_NUT.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó nal_unit_type của non-VCL NAL bằng SUFFIX_SEI_NUT.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó hình ảnh được lập mã được giải mã từ các đơn vị VCL NAL dựa vào thông điệp SEI trong đơn vị non-VCL NAL.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, còn bao gồm bước suy ra TemporalId cho đơn vị non-VCL NAL như sau: $\text{TemporalId} = \text{nuh_temporal_id_plus1} - 1$.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó giá trị của nuh_temporal_id_plus1 không bằng zero.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó TemporalId của các đơn vị

VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, còn bao gồm bước: nhận, bởi bộ giải mã, dòng bit thứ hai bao gồm một hoặc nhiều đơn vị VCL NAL thứ hai và một đơn vị non-VCL NAL thứ hai, trong đó TemporalId cho đơn vị non-VCL NAL thứ hai không bằng TemporalId của AU thứ hai chứa đơn vị non-VCL NAL thứ hai khi nal_unit_type của non-VCL NAL thứ hai là thông điệp SEI; và để đáp lại bước nhận nêu trên, thực hiện một số biện pháp hiệu chỉnh khác để bảo đảm rằng dòng bit thích hợp tương ứng với dòng bit thứ hai được nhận trước khi giải mã hình ảnh được lập mã từ các đơn vị VCL NAL thứ hai.

Theo một phương án, sáng chế đề xuất phương pháp được thực hiện bởi bộ mã hóa, phương pháp này bao gồm các bước: mã hóa, bởi bộ xử lý của bộ mã hóa, hình ảnh được lập mã trong một hoặc nhiều đơn vị VCL NAL trong dòng bit; mã hóa thành dòng bit, bởi bộ xử lý, đơn vị non-VCL NAL sao cho nuh_temporal_id_plus1 cho đơn vị non-VCL NAL được ràng buộc bằng nuh_temporal_id_plus1 của AU chứa đơn vị non-VCL NAL khi nal_unit_type của non-VCL NAL là thông điệp SEI; thực hiện, bởi bộ xử lý, tập hợp các phép kiểm tra sự tương thích dòng bit trên dòng bit dựa vào thông điệp SEI; và lưu trữ, bởi bộ nhớ được nối với bộ xử lý, dòng bit để truyền thông về phía bộ giải mã.

Một chuỗi video có thể bao gồm nhiều hình ảnh. Để bảo đảm các hình ảnh được hiển thị theo thứ tự đúng, các hệ thống lập mã video có thể gán các hình ảnh cho TemporalId. Một số hệ thống lập mã video sử dụng các lớp của các hình ảnh, trong đó mỗi lớp về cơ bản bao gồm cùng video của các độ phân giải, các kích thước hình ảnh, các tốc độ khung hình v.v. khác nhau. Các hình ảnh trong các lớp khác nhau có thể được hiển thị theo cách luân phiên, tùy thuộc vào các điều kiện ở bộ giải mã. Do đó, các hình ảnh trong các lớp khác nhau được định vị ở cùng điểm trong chuỗi video chia sẻ cùng TemporalId. Ngoài ra, các hình ảnh trong các lớp khác nhau chia sẻ cùng TemporalId tạo ra AU. Ví dụ, bộ giải mã có thể hiển thị một hình ảnh được chọn từ một lớp ở mỗi AU để hiển thị một chuỗi video. Một số hệ thống lập mã video sử dụng các thông điệp SEI. Thông điệp SEI chứa thông tin không cần thiết cho quy trình giải mã để xác định

các giá trị của các mẫu trong các hình ảnh được giải mã. Ví dụ, các thông điệp SEI có thể chứa các tham số được sử dụng bởi bộ giải mã tham chiếu giả định (HRD) hoạt động ở bộ mã hóa để kiểm tra dòng bit về sự tương thích với các chuẩn. Ngoài ra, các hệ thống lập mã video có thể lập mã chuỗi video thành dòng bit dưới dạng các lớp của các hình ảnh. Các thông điệp SEI có thể có liên quan đến các hình ảnh khác nhau và/hoặc các tổ hợp khác nhau của các lớp. Do đó, việc bảo đảm rằng thông điệp SEI thích hợp được liên kết với các hình ảnh/các lớp thích hợp trở thành thách thức trong các dòng bit đa lớp phức tạp. Trong trường hợp thông điệp SEI không được liên kết với lớp/hình ảnh đúng, HRD có thể không thể kiểm tra lớp/hình ảnh về sự tương thích theo cách đúng đắn. Việc này có thể dẫn đến các sai số mã hóa.

Ví dụ này bao gồm cơ chế để liên kết đúng các thông điệp SEI với các hình ảnh/các lớp tương ứng. Các dòng bit đa lớp có thể tổ chức các hình ảnh và các tham số có liên quan thành các AU. AU là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian đầu ra. Thông điệp SEI có thể được định vị trong cùng AU dưới dạng hình ảnh thứ nhất được liên kết với thông điệp SEI. Ngoài ra, thông điệp SEI còn được gán TemporalId. TemporalId là mã định danh biểu thị vị trí tương đối của đơn vị NAL trong chuỗi video. TemporalId của thông điệp SEI được ràng buộc bằng TemporalId của AU chứa thông điệp SEI. Nói cách khác, các hình ảnh có trong các đơn vị VCL NAL và các tham số có trong các đơn vị non-VCL NAL. Khi đơn vị non-VCL NAL là đơn vị SEI NAL chứa thông điệp SEI, TemporalId của đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL này. Phương pháp này bảo đảm rằng các thông điệp SEI được liên kết đúng với các hình ảnh tương ứng trong các AU. Do đó, có thể tránh được các lỗi khác nhau. Kết quả là, chức năng của bộ mã hóa và bộ giải mã được nâng cao. Ngoài ra, hiệu quả lập mã có thể tăng lên, giảm việc sử dụng tài nguyên báo hiệu bộ xử lý, bộ nhớ và/hoặc mạng ở cả bộ mã hóa lẫn bộ giải mã.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó nal_unit_type của non-VCL NAL bằng PREFIX_SEI_NUT.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó `nal_unit_type` của non-VCL NAL bằng `SUFFIX_SEI_NUT`.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó giá trị của `nuh_temporal_id_plus1` không bằng zero.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó `nuh_temporal_id_plus1` của các đơn vị VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU.

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

Theo một phương án, sáng chế đề xuất phương tiện không tạm thời đọc được bởi máy tính bao gồm sản phẩm chương trình máy tính để sử dụng bởi thiết bị lập mã video, sản phẩm chương trình máy tính này bao gồm các lệnh thực thi được bởi máy tính được lưu trữ trên phương tiện không tạm thời đọc được bởi máy tính sao cho khi được thực thi bởi bộ xử lý sẽ khiến thiết bị lập mã video thực hiện phương pháp theo khía cạnh bất kỳ trong số các khía cạnh nêu trên.

Theo một phương án, sáng chế đề xuất bộ giải mã bao gồm: phương tiện nhận để nhận dòng bit bao gồm hình ảnh được lập mã trong một hoặc nhiều đơn vị VCL NAL và đơn vị non-VCL NAL, trong đó `TemporalId` cho đơn vị non-VCL NAL được ràng buộc bằng `TemporalId` của AU chứa đơn vị non-VCL NAL khi `nal_unit_type` của non-VCL NAL là thông điệp SEI; phương tiện giải mã để giải mã hình ảnh được lập mã từ các đơn vị VCL NAL để tạo ra hình ảnh được giải mã; và phương tiện chuyển tiếp để chuyển tiếp hình ảnh được giải mã để hiển thị dưới dạng một phần của chuỗi video được giải mã.

Một chuỗi video có thể bao gồm nhiều hình ảnh. Để bảo đảm các hình ảnh được hiển thị theo thứ tự đúng, các hệ thống lập mã video có thể gán các hình ảnh cho

TemporalId. Một số hệ thống lập mã video sử dụng các lớp của các hình ảnh, trong đó mỗi lớp về cơ bản bao gồm cùng video của các độ phân giải, các kích thước hình ảnh, các tốc độ khung hình v.v. khác nhau. Các hình ảnh trong các lớp khác nhau có thể được hiển thị theo cách luân phiên, tùy thuộc vào các điều kiện ở bộ giải mã. Do đó, các hình ảnh trong các lớp khác nhau được định vị ở cùng điểm trong chuỗi video chia sẻ cùng TemporalId. Ngoài ra, các hình ảnh trong các lớp khác nhau chia sẻ cùng TemporalId tạo ra AU. Ví dụ, bộ giải mã có thể hiển thị một hình ảnh được chọn từ một lớp ở mỗi AU để hiển thị một chuỗi video. Một số hệ thống lập mã video sử dụng các thông điệp SEI. Thông điệp SEI chứa thông tin không cần thiết cho quy trình giải mã để xác định các giá trị của các mẫu trong các hình ảnh được giải mã. Ví dụ, các thông điệp SEI có thể chứa các tham số được sử dụng bởi bộ giải mã tham chiếu giả định (HRD) hoạt động ở bộ mã hóa để kiểm tra dòng bit về sự tương thích với các chuẩn. Ngoài ra, các hệ thống lập mã video có thể lập mã chuỗi video thành dòng bit dưới dạng các lớp của các hình ảnh. Các thông điệp SEI có thể có liên quan đến các hình ảnh khác nhau và/hoặc các tổ hợp khác nhau của các lớp. Do đó, việc bảo đảm rằng thông điệp SEI thích hợp được liên kết với các hình ảnh/các lớp thích hợp trở thành thách thức trong các dòng bit đa lớp phức tạp. Trong trường hợp thông điệp SEI không được liên kết với lớp/hình ảnh đúng, HRD có thể không thể kiểm tra lớp/hình ảnh về sự tương thích theo cách đúng đắn. Việc này có thể dẫn đến các sai số mã hóa.

Ví dụ này bao gồm cơ chế để liên kết đúng các thông điệp SEI với các hình ảnh/các lớp tương ứng. Các dòng bit đa lớp có thể tổ chức các hình ảnh và các tham số có liên quan thành các AU. AU là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian đầu ra. Thông điệp SEI có thể được định vị trong cùng AU dưới dạng hình ảnh thứ nhất được liên kết với thông điệp SEI. Ngoài ra, thông điệp SEI được gán TemporalId. TemporalId là mã định danh biểu thị vị trí tương đối của đơn vị NAL trong chuỗi video. TemporalId của thông điệp SEI được ràng buộc bằng TemporalId của AU chứa thông điệp SEI. Nói cách khác, các hình ảnh có trong các đơn vị VCL NAL và các tham số có trong các đơn vị non-VCL NAL. Khi đơn vị non-VCL NAL là đơn vị SEI NAL chứa thông điệp SEI, TemporalId của đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL này. Phương pháp này bảo đảm rằng các thông điệp SEI được liên kết đúng với các hình ảnh tương ứng trong các AU. Do đó, có thể tránh được các lỗi khác nhau. Kết quả là, chức

năng của bộ mã hóa và bộ giải mã được nâng cao. Ngoài ra, hiệu quả lập mã có thể tăng lên, giảm việc sử dụng tài nguyên báo hiệu bộ xử lý, bộ nhớ và/hoặc mạng ở cả bộ mã hóa lẫn bộ giải mã.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó bộ giải mã còn được tạo cấu hình để thực hiện phương pháp theo khía cạnh bất kỳ trong số các khía cạnh nêu trên.

Theo một phương án, sáng chế đề xuất bộ mã hóa bao gồm: phương tiện mã hóa để: mã hóa hình ảnh được lập mã trong một hoặc nhiều đơn vị VCL NAL trong dòng bit; và mã hóa thành dòng bit đơn vị non-VCL NAL sao cho TemporalId cho đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL khi nal_unit_type của non-VCL NAL là thông điệp SEI bổ sung; phương tiện HRD để thực hiện tập hợp các phép kiểm tra sự tương thích dòng bit trên dòng bit dựa vào thông điệp SEI; và phương tiện lưu trữ để lưu trữ dòng bit để truyền thông về phía bộ giải mã.

Một chuỗi video có thể bao gồm nhiều hình ảnh. Để bảo đảm các hình ảnh được hiển thị theo thứ tự đúng, các hệ thống lập mã video có thể gán các hình ảnh cho TemporalId. Một số hệ thống lập mã video sử dụng các lớp của các hình ảnh, trong đó mỗi lớp về cơ bản bao gồm cùng video của các độ phân giải, các kích thước hình ảnh, các tốc độ khung hình v.v. khác nhau. Các hình ảnh trong các lớp khác nhau có thể được hiển thị theo cách luân phiên, tùy thuộc vào các điều kiện ở bộ giải mã. Do đó, các hình ảnh trong các lớp khác nhau được định vị ở cùng điểm trong chuỗi video chia sẻ cùng TemporalId. Ngoài ra, các hình ảnh trong các lớp khác nhau chia sẻ cùng TemporalId tạo ra AU. Ví dụ, bộ giải mã có thể hiển thị một hình ảnh được chọn từ một lớp ở mỗi AU để hiển thị một chuỗi video. Một số hệ thống lập mã video sử dụng các thông điệp SEI. Thông điệp SEI chứa thông tin không cần thiết cho quy trình giải mã để xác định các giá trị của các mẫu trong các hình ảnh được giải mã. Ví dụ, các thông điệp SEI có thể chứa các tham số được sử dụng bởi bộ giải mã tham chiếu giả định (HRD) hoạt động ở bộ mã hóa để kiểm tra dòng bit về sự tương thích với các chuẩn. Ngoài ra, các hệ thống lập mã video có thể lập mã chuỗi video thành dòng bit dưới dạng các lớp của các hình ảnh. Các thông điệp SEI có thể có liên quan đến các hình ảnh khác nhau và/hoặc các tổ hợp khác nhau của các lớp. Do đó, việc bảo đảm rằng thông điệp SEI thích hợp được liên kết với các hình ảnh/các lớp thích hợp trở thành thách thức trong các dòng bit đa

lớp phức tạp. Trong trường hợp thông điệp SEI không được liên kết với lớp/hình ảnh đúng, HRD có thể không thể kiểm tra lớp/hình ảnh về sự tương thích theo cách đúng đắn. Việc này có thể dẫn đến các sai số mã hóa.

Ví dụ này bao gồm cơ chế để liên kết đúng các thông điệp SEI với các hình ảnh/các lớp tương ứng. Các dòng bit đa lớp có thể tổ chức các hình ảnh và các tham số có liên quan thành các AU. AU là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian đầu ra. Thông điệp SEI có thể được định vị trong cùng AU dưới dạng hình ảnh thứ nhất được liên kết với thông điệp SEI. Ngoài ra, thông điệp SEI được gán TemporalId. TemporalId là mã định danh biểu thị vị trí tương đối của đơn vị NAL trong chuỗi video. TemporalId của thông điệp SEI được ràng buộc bằng TemporalId của AU chứa thông điệp SEI. Nói cách khác, các hình ảnh có trong các đơn vị VCL NAL và các tham số có trong các đơn vị non-VCL NAL. Khi đơn vị non-VCL NAL là đơn vị SEI NAL chứa thông điệp SEI, TemporalId của đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL này. Phương pháp này bảo đảm rằng các thông điệp SEI được liên kết đúng với các hình ảnh tương ứng trong các AU. Do đó, có thể tránh được các lỗi khác nhau. Kết quả là, chức năng của bộ mã hóa và bộ giải mã được nâng cao. Ngoài ra, hiệu quả lập mã có thể tăng lên, giảm việc sử dụng tài nguyên báo hiệu bộ xử lý, bộ nhớ và/hoặc mạng ở cả bộ mã hóa lẫn bộ giải mã.

Một cách tùy chọn, theo khía cạnh bất kỳ trong số các khía cạnh nêu trên, phương án thực hiện khác của khía cạnh này được đề xuất, trong đó bộ mã hóa còn được tạo cấu trúc để thực hiện phương pháp theo khía cạnh bất kỳ trong số các khía cạnh nêu trên.

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

Các dấu hiệu này và khác nữa sẽ được hiểu rõ ràng hơn nhờ phần mô tả chi tiết dưới đây cùng với các hình vẽ và yêu cầu bảo hộ kèm theo.

Mô tả vắn tắt các hình vẽ

Để hiểu sáng chế rõ hơn, bây giờ thực hiện tham chiếu đến phần mô tả vắn tắt sau đây, cùng với các hình vẽ kèm theo và phần mô tả chi tiết sáng chế, trong đó các số chỉ dẫn giống nhau biểu diễn các phần giống nhau.

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

FIG.2 là sơ đồ giản lược của hệ thống lập mã và giải mã làm ví dụ (bộ lập-giải mã) để lập mã video;

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

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

FIG.5 là sơ đồ giản lược minh họa bộ giải mã tham chiếu giả định (HRD) làm ví dụ;

FIG.6 là sơ đồ giản lược minh họa chuỗi video đa lớp làm ví dụ;

FIG.7 là sơ đồ giản lược minh họa dòng bit làm ví dụ;

FIG.8 là sơ đồ giản lược của thiết bị lập mã video làm ví dụ;

FIG.9 là lưu đồ của phương pháp làm ví dụ để mã hóa chuỗi video thành dòng bit bằng cách ràng buộc các mã định danh theo thời gian (temporal identifier, TemporalId) cho các thông điệp thông tin tăng cường bổ sung (supplemental enhancement information, SEI) trong dòng bit;

FIG.10 là lưu đồ của phương pháp làm ví dụ để giải mã chuỗi video từ một dòng bit trong đó các TemporalId cho các thông điệp SEI trong dòng bit được ràng buộc;

FIG.11 là sơ đồ giản lược của một hệ thống làm ví dụ để lập mã chuỗi video sử dụng dòng bit trong đó các TemporalId cho các thông điệp SEI trong dòng bit được ràng buộc.

Mô tả chi tiết sáng chế

Cần hiểu ngay từ đầu rằng mặc dù phương án thực hiện minh họa của một hoặc nhiều phương án được đề xuất dưới đây, nhưng các hệ thống và/hoặc các phương pháp được mô tả có thể được thực hiện bằng cách sử dụng một số kỹ thuật bất kỳ, bất kể đã biết hay đang tồn tại. Sáng chế không bị giới hạn ở các phương án thực hiện minh họa, các hình vẽ, và các kỹ thuật được minh họa dưới đây, bao gồm các thiết kế ví dụ và các

phương án thực hiện được minh họa và được mô tả ở đây, mà có thể được cải biến trong phạm vi của yêu cầu bảo hộ kèm theo cùng với toàn bộ phạm vi tương đương của yêu cầu bảo hộ này.

Các thuật ngữ sau được định nghĩa như dưới đây trừ khi được sử dụng theo ngữ cảnh ngược lại trong bản mô tả này. Cụ thể là, các định nghĩa sau nhằm giúp là rõ ràng sáng chế hơn. Tuy nhiên, các thuật ngữ có thể được mô tả khác nhau trong các ngữ cảnh khác nhau. Do đó, các định nghĩa dưới đây nên được xem là bổ sung và không nên được xem là giới hạn các định nghĩa mô tả bất kỳ khác được đưa ra cho các thuật ngữ này trong bản mô tả.

Dòng bit là chuỗi các bit bao gồm dữ liệu video được nén để truyền dẫn giữa bộ mã hóa và bộ giải mã. Bộ mã hóa là thiết bị được tạo cấu hình để sử dụng các quy trình mã hóa để nén dữ liệu video thành dòng bit. Bộ giải mã là thiết bị được tạo cấu hình để sử dụng các quy trình giải mã để tái dựng dữ liệu video từ một dòng bit để hiển thị. Hình ảnh là mảng gồm các mẫu độ chói (luma) và/hoặc mảng gồm các mẫu sắc độ (chroma) tạo thành khung hoặc trường của hình ảnh này. Lát (slice) là số lượng dạng số nguyên của các ô (tile) hoàn chỉnh hoặc số lượng dạng số nguyên của các hàng đơn vị cây lập mã (coding tree unit, CTU) hoàn chỉnh liên tiếp (ví dụ, trong ô) của hình ảnh mà được chứa riêng trong một đơn vị lớp trừu tượng hóa mạng (network abstraction layer, NAL). Hình ảnh đang được mã hóa hoặc được giải mã có thể được gọi là hình ảnh hiện thời để rõ ràng cho việc thảo luận. Hình ảnh được lập mã là biểu diễn được lập mã của hình ảnh bao gồm các đơn vị VCL NAL (đơn vị NAL lớp lập mã video (video coding layer (VCL) NAL unit)) có giá trị cụ thể là mã định danh lớp phân đầu đơn vị NAL (nuh_layer_id) bên trong đơn vị truy cập (access unit, AU) và chứa tất cả các đơn vị cây lập mã (coding tree unit, CTU) của hình ảnh. Hình ảnh được giải mã là hình ảnh được tạo ra bằng cách áp dụng quy trình giải mã cho hình ảnh được lập mã.

AU là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian để kết xuất từ bộ đệm hình ảnh được giải mã (coded picture buffer, DPB). Đơn vị NAL là cấu trúc cú pháp chứa dữ liệu dưới dạng tải trọng chuỗi byte thô (Raw Byte Sequence Payload, RBSP), chỉ báo về kiểu dữ liệu, và được đặt rải rác các byte chống giả lập theo ý muốn. Đơn vị VCL NAL là đơn vị NAL được lập mã để chứa dữ liệu video, chẳng hạn như lát được lập mã của hình ảnh. Đơn vị non-VCL NAL (đơn

vị NAL phi VCL) là đơn vị NAL chứa dữ liệu không phải video như cú pháp và/hoặc các tham số hỗ trợ việc giải mã dữ liệu video, thực hiện kiểm tra tính tương thích, hoặc các hoạt động khác. Kiểu đơn vị NAL (`nal_unit_type`) là phần tử cú pháp được chứa trong đơn vị NAL biểu thị kiểu của dữ liệu được chứa trong đơn vị NAL. Lớp là tập hợp các đơn vị VCL NAL chia sẻ một đặc trưng xác định (ví dụ độ phân giải, tốc độ khung hình, kích thước hình ảnh chung, v.v.) như được biểu thị bởi ID lớp và các đơn vị non-VCL NAL có liên quan. Mã định danh lớp phần đầu đơn vị NAL (`nuh_layer_id`) là phần tử cú pháp xác định mã định danh của lớp bao gồm đơn vị NAL. Mã định danh theo thời gian (`TemporalId`) là mã định danh được suy ra biểu thị vị trí tương đối của đơn vị NAL trong chuỗi video. Mã định danh theo thời gian phần đầu đơn vị NAL cộng một (`nuh_temporal_id_plus1`) là mã định danh được báo hiệu biểu thị vị trí tương đối của đơn vị NAL trong chuỗi video.

Bộ giải mã tham chiếu giả định (hypothetical reference decoder, HRD) là mô hình bộ giải mã hoạt động trên bộ mã hóa kiểm tra tính biến thiên của các dòng bit được tạo ra bởi quy trình mã hóa để xác nhận sự tương thích với các ràng buộc xác định. Kiểm tra tương thích dòng bit là phép kiểm tra để xác định liệu dòng bit được mã hóa có phù hợp với một chuẩn hay không, như lập mã video đa năng (Versatile Video Coding, VVC) chẳng hạn. Các tham số HRD là các phần tử cú pháp khởi tạo và/hoặc xác định các điều kiện hoạt động của HRD. Các tham số HRD có thể có trong các thông điệp thông tin tăng cường bổ sung (supplemental enhancement information, SEI) và/hoặc trong tập hợp tham số video (video parameter set, VPS). Thông điệp SEI là cấu trúc cú pháp có ngữ nghĩa xác định chuyển tải thông tin không cần thiết cho quy trình giải mã để xác định các giá trị của các mẫu trong các hình ảnh được giải mã. Đơn vị SEI NAL là đơn vị NAL chứa một hoặc nhiều thông điệp SEI. Đơn vị SEI NAL cụ thể có thể được gọi là đơn vị SEI NAL hiện thời. Thông điệp SEI lồng định tỷ lệ được là thông điệp chứa nhiều thông điệp SEI tương ứng với một hoặc nhiều tập hợp lớp đầu ra (output layer set, OLS) hoặc một hoặc nhiều lớp. Thông điệp SEI chu kỳ đệm (buffering period (BP) SEI message) là thông điệp SEI chứa các tham số HRD để khởi tạo HRD để quản lý bộ đệm hình ảnh được lập mã (coded picture buffer, CPB). Thông điệp SEI định thời hình ảnh (picture timing (PT) SEI message) là thông điệp SEI chứa các tham số HRD để quản lý thông tin phân phối cho các AU ở CPB và/hoặc bộ đệm hình ảnh được giải mã (decoded picture buffer, DPB). Thông điệp SEI thông tin đơn vị giải mã (decoding unit

information (DUI) SEI message) là thông điệp SEI chứa các tham số HRD để quản lý thông tin phân phối cho các DU ở CPB và/hoặc DPB. Thông điệp SEI lồng định tỷ lệ được là tập hợp gồm các thông điệp SEI được lồng-định tỷ lệ được. Thông điệp SEI được lồng-định tỷ lệ được là thông điệp SEI được lồng bên trong thông điệp SEI lồng định tỷ lệ được. Thông điệp SEI tiền tố là thông điệp SEI áp dụng cho một hoặc nhiều đơn vị NAL sao đó. Thông điệp SEI hậu tố là thông điệp SEI áp dụng cho một hoặc nhiều đơn vị NAL đứng trước.

Tập hợp tham số hình ảnh (picture parameter set, PPS) là cấu trúc cú pháp chứa các phần tử cú pháp áp dụng cho toàn bộ các hình ảnh được lập mã như được xác định bởi phần tử cú pháp được tìm thấy trong mỗi phần đầu hình ảnh. Phần đầu hình ảnh là cấu trúc cú pháp chứa các phần tử cú pháp áp dụng cho tất cả các lát của hình ảnh được lập mã. Phần đầu lát là một phần của lát được lập mã chứa các phần tử dữ liệu liên quan đến tất cả các ô hoặc các hàng CTU trong ô được biểu diễn trong lát. Chuỗi video được lập mã là tập hợp của một hoặc nhiều hình ảnh được lập mã. Chuỗi video được giải mã là tập hợp của một hoặc nhiều hình ảnh được giải mã.

Các cụm từ viết tắt sau được sử dụng dưới đây, đơn vị truy cập (access unit, AU), khối cây lập mã (Coding Tree Block, CTB), đơn vị cây lập mã (Coding Tree Unit, CTU), đơn vị lập mã (Coding Unit, CU), chuỗi video lớp được lập mã (Coded Layer Video Sequence, CLVS), phần bắt đầu chuỗi video lớp được lập mã (Coded Layer Video Sequence Start, CLVSS), chuỗi video được lập mã (Coded Video Sequence, CVS), phần bắt đầu chuỗi video được lập mã (Coded Video Sequence Start, CVSS), đội ngũ chung các chuyên gia về video (Joint Video Experts Team, JVET), bộ giải mã tham chiếu giả định (Hypothetical Reference Decoder, HRD), tập hợp ô ràng buộc chuyển động (Motion Constrained Tile Set, MCTS), đơn vị chuyển tải lớn nhất (Maximum Transfer Unit, MTU), lớp trừu tượng hóa mạng (Network Abstraction Layer, NAL), tập hợp lớp đầu ra (Output Layer Set, OLS), bộ đếm thứ tự hình ảnh (Picture Order Count, POC), điểm truy cập ngẫu nhiên (Random Access Point, RAP), tải trọng chuỗi byte thô (Raw Byte Sequence Payload, RBSP), tập hợp tham số trình tự (Sequence Parameter Set, SPS), tập hợp tham số video (video parameter set, VPS), lập mã video đa năng (Versatile Video Coding, VVC).

Nhiều kỹ thuật nén video có thể được sử dụng để giảm kích thước của các tệp tin video với tổn thất dữ liệu nhỏ nhất. Ví dụ, các kỹ thuật nén video có thể bao gồm thực hiện phép dự đoán theo không gian (ví dụ nội-hình ảnh) và/hoặc phép dự đoán theo thời gian (ví dụ liên-hình ảnh) để giảm hoặc loại bỏ sự dư thừa dữ liệu trong các chuỗi video. Đối với việc lập mã video dựa vào khối, lát video (ví dụ hình ảnh video hoặc một phần của hình ảnh video) có thể được phân vùng thành các khối video, cũng có thể được gọi là các khối cây, các khối cây lập mã (CTB), các đơn vị cây lập mã (CTU), các đơn vị lập mã (CU), và/hoặc các nút lập mã. Các khối video trong lát được lập mã nội ảnh (I) của hình ảnh được lập mã sử dụng phép dự đoán theo không gian so với các mẫu tham chiếu trong các khối lân cận trong cùng hình ảnh. Các khối video trong lát dự đoán một chiều được lập mã liên ảnh (P) hoặc lát dự đoán hai chiều (B) của hình ảnh có thể được lập mã bằng cách sử dụng phép dự đoán theo không gian so với các mẫu tham chiếu trong các khối lân cận trong cùng hình ảnh hoặc phép dự đoán theo thời gian so với các mẫu tham chiếu trong các hình ảnh tham chiếu khác. Các hình ảnh có thể được gọi là các khung và/hoặc các ảnh, và các hình ảnh tham chiếu có thể được gọi là các khung tham chiếu và/hoặc các ảnh tham chiếu. Kết quả của phép dự đoán theo không gian hoặc theo thời gian là khối dự đoán biểu diễn khối ảnh. Dữ liệu dư biểu diễn các khác biệt điểm ảnh giữa khối ảnh gốc và khối dự đoán. Do đó, khối được lập mã liên ảnh được mã hóa theo vector chuyển động trở vào khối của các mẫu tham chiếu tạo ra khối dự đoán và dữ liệu dư biểu thị khác biệt giữa khối được lập mã và khối dự đoán. Khối được lập mã nội ảnh được mã hóa theo chế độ lập mã nội ảnh và dữ liệu dư. Để nén hơn nữa, dữ liệu dư có thể được biến đổi từ miền điểm ảnh sang miền biến đổi. Kết quả của việc này là các hệ số biến đổi dư, mà có thể được lượng tử hóa. Các hệ số biến đổi được lượng tử hóa có thể được sắp xếp ban đầu trong mảng hai chiều. Các hệ số biến đổi được lượng tử hóa có thể được quét để tạo ra vector một chiều của các hệ số biến đổi. Lập mã entropi có thể được áp dụng để thu được nén thêm. Các kỹ thuật nén video này được trình bày chi tiết hơn dưới đây.

Để bảo đảm video được mã hóa có thể được giải mã chính xác, video được mã hóa và được giải mã theo các chuẩn lập mã video tương ứng. Các chuẩn lập mã video bao gồm chuẩn ITU-T H.261 (International Telecommunication Union (ITU) Standardization Sector: bộ phận chuẩn hóa hiệp hội viễn thông quốc tế), ISO/IEC MPEG-1 Phần 2 (International Organization for Standardization/International

Electrotechnical Commission (ISO/IEC) Motion Picture Experts Group (MPEG) - nhóm chuyên gia về hình ảnh di động của tổ chức quốc tế về chuẩn hóa/ủy ban kỹ thuật điện quốc tế), ITU-T H.262 hoặc ISO/IEC MPEG-2 Phần 2, ITU-T H.263, ISO/IEC MPEG-4 Phần 2, lập mã video tiên tiến (Advanced Video Coding, AVC), còn được gọi là ITU-T H.264 hoặc ISO/IEC MPEG-4 Phần 10, và lập mã video hiệu suất cao (High Efficiency Video Coding, HEVC), còn được gọi là ITU-T H.265 hoặc MPEG-H Phần 2. AVC bao gồm các mở rộng như lập mã video phân cấp được (Scalable Video Coding, SVC), lập mã video đa điểm (Multiview Video Coding, MVC) và lập mã video đa điểm cộng độ sâu (MVC+D), và AVC ba chiều (3D) (3D-AVC). HEVC bao gồm các mở rộng như HEVC phân cấp được (SHVC), HEVC đa điểm (MV-HEVC), và HEVC 3D (3D-HEVC). Đội ngũ chung các chuyên gia về video (joint video experts team, JVET) của ITU-T và ISO/IEC đã bắt đầu phát triển chuẩn lập mã video được gọi là lập mã video đa năng (Versatile Video, VVC). VVC có trong bản dự thảo làm việc (Working Draft, WD), bao gồm JVET-O2001-v14.

Chuỗi video có thể bao gồm nhiều hình ảnh. Để bảo đảm các hình ảnh được hiển thị theo thứ tự đúng, các hệ thống lập mã video có thể gán các hình ảnh cho mã định danh theo thời gian (TemporalId). Một số hệ thống lập mã video sử dụng các lớp của các hình ảnh, trong đó về cơ bản, mỗi lớp bao gồm cùng video của các độ phân giải, các kích thước hình ảnh, các tốc độ khung hình v.v. khác nhau. Các hình ảnh trong các lớp khác nhau có thể được hiển thị theo cách luân phiên, tùy thuộc vào các điều kiện ở bộ giải mã. Do đó, các hình ảnh trong các lớp khác nhau được định vị ở cùng điểm trong chuỗi video chia sẻ cùng TemporalId. Ngoài ra, các hình ảnh trong các lớp khác nhau chia sẻ cùng TemporalId tạo ra đơn vị truy cập (access unit, AU). Ví dụ, bộ giải mã có thể hiển thị một hình ảnh được chọn từ một lớp ở mỗi AU để hiển thị một chuỗi video.

Một số hệ thống lập mã video sử dụng các thông điệp SEI. Thông điệp SEI chứa thông tin không cần thiết cho quy trình giải mã để xác định các giá trị của các mẫu trong các hình ảnh được giải mã. Ví dụ, các thông điệp SEI có thể chứa các tham số được sử dụng bởi HRD hoạt động ở bộ mã hóa để kiểm tra dòng bit về sự tương thích với các chuẩn. Ngoài ra, các hệ thống lập mã video cũng có thể lập mã chuỗi video thành dòng bit dưới dạng các lớp của các hình ảnh. Các thông điệp SEI có thể có liên quan đến các hình ảnh khác nhau và/hoặc các tổ hợp khác nhau của các lớp. Do đó, việc bảo đảm rằng

thông điệp SEI thích hợp được liên kết với các hình ảnh/các lớp thích hợp trở thành thách thức trong các dòng bit đa lớp phức tạp. Trong trường hợp thông điệp SEI không được liên kết với lớp/hình ảnh đúng, HRD có thể không thể kiểm tra lớp/hình ảnh về sự tương thích theo cách đúng đắn. Việc này có thể dẫn đến các sai số mã hóa.

Cơ chế được mô tả ở đây là cơ chế để liên kết đúng các thông điệp SEI với các hình ảnh/các lớp tương ứng. Các dòng bit đa lớp có thể tổ chức các hình ảnh và các tham số có liên quan thành các AU. AU là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian đầu ra. Thông điệp SEI có thể được định vị trong cùng AU khi hình ảnh thứ nhất được liên kết với thông điệp SEI. Ngoài ra, thông điệp SEI được gán TemporalId. TemporalId là mã định danh biểu thị vị trí tương đối của đơn vị lớp trừu tượng hóa mạng (network abstraction layer, NAL) trong chuỗi video. TemporalId của thông điệp SEI được ràng buộc bằng TemporalId của AU chứa thông điệp SEI. Nói cách khác, các hình ảnh có trong các đơn vị NAL lớp lập mã video (video coding layer (VCL) NAL unit) và các tham số có trong các đơn vị non-VCL NAL. Khi đơn vị non-VCL NAL là đơn vị SEI NAL chứa thông điệp SEI, thì TemporalId của đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL này. Phương pháp này bảo đảm rằng các thông điệp SEI được liên kết đúng với các hình ảnh tương ứng trong các AU. Do đó, có thể tránh được các lỗi khác nhau. Kết quả là, chức năng của bộ mã hóa và bộ giải mã được nâng cao. Ngoài ra, hiệu quả lập mã có thể tăng lên, giảm việc sử dụng tài nguyên báo hiệu bộ xử lý, bộ nhớ và/hoặc mạng ở cả bộ mã hóa lẫn bộ giải mã.

FIG.1 là lưu đồ của phương pháp hoạt động làm ví dụ 100 để lập mã tín hiệu video. Cụ thể là, tín hiệu video được mã hóa ở bộ mã hóa. Quy trình mã hóa nén tín hiệu video bằng cách sử dụng các cơ chế khác nhau để giảm kích thước tệp tin video. Kích thước tệp tin nhỏ hơn cho phép tệp tin video được nén được truyền đến người dùng, đồng thời giảm chi phí băng thông có liên quan. Sau đó, bộ giải mã giải mã tệp tin video được nén để tái dựng tín hiệu video gốc để hiển thị cho người dùng đầu cuối. Quy trình giải mã này thường đối xứng gương với quy trình mã hóa để cho phép bộ giải mã tái dựng tín hiệu video một cách nhất quán.

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

video có thể được thu lại bởi thiết bị thu video, như máy quay video chẳng hạn, và được mã hóa để hỗ trợ cho việc phát trực tiếp video. Tập tin video có thể bao gồm cả thành phần âm thanh lẫn thành phần video. Thành phần video chứa dãy các khung ảnh mà, khi được xem theo chuỗi, tạo ra cảm giác chuyển động về mặt thị giác. Các khung chứa các điểm ảnh được biểu thị về mặt ánh sáng, ở đây được gọi là các thành phần độ chói (hoặc các mẫu độ chói), và về mặt màu, được gọi là các thành phần sắc độ (hoặc các mẫu màu). Trong một số ví dụ, các khung cũng có thể chứa các giá trị độ sâu để hỗ trợ xem ba chiều.

Ở bước 103, video được phân vùng thành các khối. Việc phân vùng này bao gồm việc chia nhỏ các điểm ảnh trong mỗi khung thành các khối hình vuông và/hoặc hình chữ nhật để nén. Ví dụ, trong lập mã video hiệu suất cao (HEVC) (còn được gọi là H.265 và MPEG-H Phần 2) trước tiên, khung có thể được chia thành các đơn vị cây lập mã (CTU), mà là các khối có kích thước định trước (ví dụ, sáu mươi tư điểm ảnh nhân sáu mươi tư điểm ảnh). Các CTU chứa cả mẫu độ chói và sắc độ. Các cây lập mã có thể được sử dụng để chia các CTU thành các khối và sau đó chia nhỏ đệ quy các khối cho đến khi đạt được các cấu hình hỗ trợ việc mã hóa tiếp. Ví dụ, các thành phần độ chói của khung có thể được chia nhỏ cho đến khi các khối riêng lẻ chứa các giá trị ánh sáng tương đối đồng nhất. Ngoài ra, các thành phần sắc độ của khung cũng có thể được chia nhỏ cho đến khi các khối riêng lẻ chứa các giá trị màu sắc tương đối đồng nhất. Do đó, cơ chế phân vùng thay đổi tùy 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 ảnh và/hoặc dự đoán nội ảnh có thể được sử dụng. Dự đoán liên ảnh được thiết kế để tận dụng việc đối tượng trong một cảnh chung có xu hướng xuất hiện trong các khung liên tiếp. Do đó, khối mô tả đối tượng trong khung tham chiếu không nhất thiết được mô tả lặp lại trong các khung liên kề. Cụ thể là, đối tượng, như một cái bàn, có thể giữ nguyên ở vị trí không đổi qua nhiều khung. Do đó, cái bàn này được mô tả một lần và các khung liên kề có thể tham chiếu ngược lại khung tham chiếu. Cơ chế so khớp hình mẫu có thể được sử dụng để so khớp các đối tượng qua nhiều khung. Ngoài ra, các đối tượng di chuyển có thể được biểu diễn qua nhiều khung, ví dụ do chuyển động của đối tượng hoặc chuyển động của camera. Trong một ví dụ cụ thể, video có thể thể hiện một xe ô tô di chuyển ngang màn hình qua nhiều

khung hình. Các vector chuyển động có thể được sử dụng để mô tả chuyển động này. Vector chuyển động là vector hai chiều cho biết độ lệch từ các tọa độ của đối tượng trong khung hình đến các tọa độ của đối tượng này trong khung tham chiếu. Do đó, dự đoán liên ảnh có thể mã hóa khối ảnh trong khung hình hiện thời dưới dạng tập hợp các vector chuyển động biểu thị độ lệch so với khối tương ứng trong khung tham chiếu.

Dự đoán nội ảnh mã hóa các khối trong một khung chung. Dự đoán nội ảnh tận dụng việc các thành phần độ chói và sắc độ có xu hướng phân cụm trong một khung. Ví dụ, một mảnh màu xanh lá cây ở một phần của cây có xu hướng được bố trí liền kề các mảnh màu xanh lá cây tương tự. Dự đoán nội ảnh sử dụng nhiều chế độ dự đoán có hướng (ví dụ ba mươi ba chế độ trong HEVC), chế độ phẳng, và chế độ trung bình (direct current, DC). Các chế độ có hướng biểu thị rằng khối hiện thời là tương tự/giống các mẫu của khối lân cận theo hướng tương ứng. Chế độ phẳng biểu thị rằng dãy khối dọc theo hàng/cột (ví dụ mặt phẳng) có thể được nội suy dựa vào các khối lân cận ở các mép của hàng. Chế độ phẳng có tác dụng biểu thị sự chuyển tiếp mượt của ánh sáng/màu qua hàng/cột bằng cách sử dụng hệ số góc hầu như không đổi khi thay đổi các giá trị. Chế độ DC được sử dụng để làm mượt đường bao và biểu thị rằng khối là tương tự/giống giá trị trung bình được liên kết với các mẫu của tất cả các khối lân cận được liên kết với các hướng góc của các chế độ dự đoán có hướng. Do đó, các khối dự đoán nội ảnh có thể biểu diễn các khối ảnh dưới dạng các giá trị chế độ dự đoán tương quan khác nhau thay vì các giá trị thực. Ngoài ra, các khối dự đoán liên ảnh có thể biểu diễn các khối ảnh dưới dạng các giá trị vector chuyển động thay vì các giá trị thực. Trong cả hai trường hợp, các khối dự đoán có thể không biểu diễn chính xác các khối ảnh trong một số trường hợp. Các chênh lệch bất kỳ được lưu trữ trong các khối dư. Các phép biến đổi có thể được áp dụng cho các khối dư để nén tệp tin hơn nữa.

Ở bước 107, các kỹ thuật lọc khác nhau có thể được áp dụng. Trong chuẩn HEVC, các bộ lọc được áp dụng theo sơ đồ lọc trong vòng lặp. Phép dự đoán dựa vào khối được trình bày ở trên có thể dẫn đến việc tạo ra các hình ảnh có khối ở bộ giải mã. Ngoài ra, sơ đồ phép dự đoán dựa vào khối có thể mã hóa một khối và sau đó tái dựng khối được mã hóa để sử dụng sau đó làm khối tham chiếu. Sơ đồ lọc trong vòng lặp áp dụng lặp đi lặp lại các bộ lọc khử nhiễu, các bộ lọc tách khối, các bộ lọc vòng lặp thích ứng, và các bộ lọc độ lệch tương thích mẫu (SAO) cho các khối/các khung. Các bộ lọc này làm giảm

các nhiễu (artifact) khối này sao cho tệp tin được mã hóa có thể được tái dựng chính xác. Ngoài ra, các bộ lọc này còn làm giảm các nhiễu trong các khối tham chiếu được tái dựng sao cho các nhiễu ít có khả năng tạo ra nhiễu bổ sung trong các khối sau đó được mã hóa dựa vào các khối tham chiếu được tái dựng.

Khi tín hiệu video đã được phân vùng, được nén, và được lọc, dữ liệu thu được được mã hóa trong dòng bit ở bước 109. Dòng bit này bao gồm dữ liệu được trình bày ở trên cũng như dữ liệu báo hiệu bất kỳ được mong muốn để hỗ trợ việc tái dựng tín hiệu video 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 cung cấp các lệnh lập mã cho bộ giải mã. Dòng bit có thể được lưu trữ trong bộ nhớ để truyền dẫn về phía bộ giải mã theo yêu cầu. Dòng bit này cũng có thể phát rộng và/hoặc phát đa hướng về phía các bộ giải mã. Việc tạo ra dòng bit là quy trình lặp đi lặp lại. Do đó, các bước 101, 103, 105, 107, và 109 có thể diễn ra liên tục và/hoặc đồng thời qua nhiều khung và khối. Thứ tự được thể hiện trên Fig.1 được biểu diễn cho rõ ràng và dễ mô tả, và không nhằm giới hạn quy trình lập mã video ở một thứ tự cụ thể.

Bộ giải mã nhận dòng bit và bắt đầu quy trình giải mã ở bước 111. Cụ thể là, bộ giải mã sử dụng sơ đồ giải mã entropi để biến 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 cho các khung ở bước 111. Việc phân vùng này có thể khớp với các kết quả của việc phân vùng khối ở bước 103. Việc mã hóa/giải mã entropi như được sử dụng trong bước 111 sẽ được mô tả. Bộ mã hóa thực hiện nhiều lựa chọn trong quy trình nén, như lựa chọn sơ đồ phân vùng khối từ một số lựa chọn khả thi dựa vào việc định vị theo không gian của các giá trị trong (các) hình ảnh đầu vào. Việc báo hiệu các lựa chọn chính xác có thể sử dụng số lượng bin lớn. Như được sử dụng ở đây, bin là giá trị nhị phân được xử lý dưới dạng biến số (ví dụ, giá trị bit có thể thay đổi tùy thuộc vào ngữ cảnh). Việc lập mã entropi cho phép bộ mã hóa loại bỏ các tùy chọn bất kỳ rõ ràng không thể làm được đối với trường hợp cụ thể, để lại tập hợp các tùy chọn được cho phép. Sau đó, mỗi tùy chọn được cho phép được gán một từ mã. Độ dài của các từ mã dựa vào số lượng tùy chọn được cho phép (ví dụ một bin cho hai tùy chọn, hai bin cho từ hai đến ba tùy chọn, v.v.). Sau đó, bộ mã hóa mã hóa từ mã cho tùy chọn được chọn. Sơ đồ này giảm kích thước của các từ mã khi các từ mã này càng lớn càng tốt để biểu thị duy nhất

sự lựa chọn từ một tập hợp con nhỏ gồm các tùy chọn được cho phép trái ngược với việc biểu thị duy nhất sự lựa chọn từ tập hợp rộng sẵn có gồm tất cả các tùy chọn khả thi. Sau đó, bộ giải mã giải mã sự lựa chọn bằng cách xác định tập hợp các tùy chọn được cho phép theo cách tương tự như bộ mã hóa. Bằng cách xác định tập hợp các tùy chọn được cho 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 việc giải mã khối. Cụ thể là, bộ giải mã sử dụng phép biến đổi ngược để tạo ra các khối dư. Sau đó, bộ giải mã sử dụng các khối dư và các khối dự đoán tương ứng để tái dựng các khối ảnh theo việc phân vùng này. Các khối dự đoán có thể bao gồm cả các khối dự đoán nội ảnh lẫn các khối dự đoán liên ảnh như được tạo ra ở bộ mã hóa ở bước 105. Sau đó, các khối ảnh được tái dựng được định vị vào các khung của tín hiệu video được tái dựng theo việc phân vùng dữ liệu được xác định ở bước 111. Cú pháp cho bước 113 có thể cũng được báo hiệu trong dòng bit thông qua việc lập mã entropi như được trình bày ở trên.

Ở bước 115, việc lọc được thực hiện trên các khung của tín hiệu video được tái dựng theo cách tương tự bước 107 ở bộ mã hóa. Ví dụ, các bộ lọc khử nhiễu, các bộ lọc tách khối, các bộ lọc vòng lặp thích ứng, và các bộ lọc SAO có thể được áp dụng cho các khung để loại bỏ các nhiễu khối. Vì các khung này được lọc, nên tín hiệu video có thể được xuất ra màn hình ở bước 117 để xem bởi người dùng đầu cuối.

FIG.2 là sơ đồ giản lược của hệ thống lập mã và giải mã (bộ lập-giải mã) 200 làm ví dụ để lập mã video. Cụ thể là, hệ thống lập-giải mã 200 cung cấp chức năng hỗ trợ việc thực hiện phương pháp hoạt động 100. Hệ thống lập-giải mã 200 được tổng quát hóa để thể hiện các thành phần được sử dụng trong cả bộ mã hóa lẫn bộ giải mã. Hệ thống lập-giải mã 200 nhận và phân vùng tín hiệu video như được trình bày đối với các bước 101 và 103 trong phương pháp hoạt động 100, tạo ra tín hiệu video được phân vùng 201. Sau đó, hệ thống lập-giải mã 200 nén tín hiệu video được phân vùng 201 thành dòng bit được lập mã khi đóng vai trò là bộ mã hóa như được trình bày đối với các bước 105, 107, và 109 trong phương pháp 100. Khi đóng vai trò là bộ giải mã, hệ thống lập-giải mã 200 tạo ra tín hiệu video đầu ra từ dòng bit như được trình bày đối với các bước 111, 113, 115, và 117 trong phương pháp hoạt động 100. Hệ thống lập-giải mã 200 bao gồm thành phần điều khiển bộ lập mã chung 211, thành phần định tỷ lệ biến đổi

và lượng tử hóa 213, thành phần ước lượng nội-hình ảnh 215, thành phần dự đoán nội-hình ảnh 217, thành phần bù chuyển động 219, thành phần ước lượng chuyển động 221, thành phần định tỷ lệ và biến đổi ngược 229, thành phần phân tích điều khiển bộ lọc 227, thành phần lọc trong vòng lặp 225, thành phần đệm hình ảnh được giải mã 223, và thành phần định dạng phần đầu và lập mã số học nhị phân thích ứng ngữ cảnh (context adaptive binary arithmetic coding, CABAC) 231. Các thành phần này được nối như được thể hiện trên hình vẽ. Trên FIG.2, các đường nét liền biểu thị sự dịch chuyển của dữ liệu được mã hóa/được giải mã trong khi các đường nét đứt biểu thị sự dịch chuyển của dữ liệu điều khiển điều khiển hoạt động của các thành phần khác. Các thành phần của hệ thống lập-giải mã 200 có thể đều có mặt trong bộ mã hóa. Bộ giải mã có thể bao gồm tập hợp con gồm các thành phần của hệ thống lập-giải mã 200. Ví dụ, bộ giải mã có thể bao gồm thành phần dự đoán nội-hình ảnh 217, thành phần bù chuyển động 219, thành phần định tỷ lệ và biến đổi ngược 229, thành phần lọc trong vòng lặp 225, và thành phần đệm hình ảnh được giải mã 223. Các thành phần này sẽ được mô tả dưới đây.

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

Thành phần điều khiển bộ lập mã chung 211 được tạo cấu hình để đưa ra các quyết định có liên quan đến việc lập mã các hình ảnh của chuỗi video thành dòng bit theo các ràng buộc ứng dụng. Ví dụ, thành phần điều khiển bộ lập mã chung 211 quản lý sự tối ưu của tốc độ bit/kích thước dòng bit so với chất lượng tái dựng. Các quyết định này có thể được đưa ra dựa vào tính khả dụng không gian lưu trữ/băng thông và các yêu cầu độ phân giải hình ảnh. Thành phần điều khiển bộ lập mã chung 211 cũng quản lý việc sử dụng bộ đệm theo tốc độ truyền để làm giảm các vấn đề tràn bộ đệm. Để quản lý các vấn đề này, thành phần điều khiển bộ lập mã chung 211 quản lý việc phân vùng, dự đoán và lọc bởi các thành phần khác. Ví dụ, thành phần điều khiển bộ lập mã chung 211 có thể tăng theo cách động độ phức tạp khi nén để tăng độ phân giải và tăng lượng sử dụng băng thông hoặc giảm độ phức tạp khi nén để giảm độ phân giải và lượng sử dụng băng thông. Do đó, thành phần điều khiển bộ lập mã chung 211 điều khiển các thành phần khác của hệ thống lập-giải mã 200 để làm cân bằng chất lượng tái dựng tín hiệu video với các vấn đề về tốc độ bit. Thành phần điều khiển bộ lập mã chung 211 tạo ra dữ liệu điều khiển, điều khiển hoạt động của các thành phần khác. Dữ liệu điều khiển này cũng được chuyển tiếp đến thành phần định dạng phần đầu và CABAC 231 để được mã hóa trong dòng bit để báo hiệu các tham số để giải mã ở bộ giải mã.

Tín hiệu video được phân vùng 201 còn được gửi đến thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 để dự đoán liên ảnh. 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 lập mã dự đoán liên ảnh cho khối video nhận được so với một hoặc nhiều khối trong một hoặc nhiều khung tham chiếu để thực hiện phép dự đoán theo thời gian. Hệ thống lập-giải mã 200 có thể thực hiện nhiều đường dẫn lập mã, ví dụ để chọn chế độ lập mã thích 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 ở mức độ cao, nhưng được minh họa riêng biệt cho các mục đích nhận thức. Việc ước lượng chuyển động, được thực hiện bởi thành phần ước lượng chuyển động 221, là quy trình tạo ra các vector chuyển động, ước lượng chuyển động cho các khối video. Ví dụ, vector chuyển động có thể biểu thị sự dịch chuyển của đối tượng được lập mã so với khối dự đoán. Khối dự đoán là khối được tìm ra phù hợp sát với khối được

lập mã, về mặt chênh lệch điểm ảnh. Khối dự đoán cũng có thể được gọi là khối tham chiếu. Chênh lệch điểm ảnh này có thể được xác định bởi tổng của chênh lệch tuyệt đối (sum of absolute difference, SAD), tổng của chênh lệch bình phương (sum of square difference, SSD), hoặc các số liệu chênh lệch khác. HEVC sử dụng một số đối tượng được lập mã bao gồm CTU, các khối cây lập mã (CTB), và các CU. Ví dụ, CTU có thể được chia thành các CTB, sau đó có thể được chia thành các CB để bao gồm trong các CU. Một CU có thể được mã hóa dưới dạng đơn vị dự đoán chứa dữ liệu dự đoán và/hoặc đơn vị biến đổi (TU) chứa dữ liệu dư được biến đổi cho CU. Thành phần ước lượng chuyển động 221 tạo ra các vector chuyển động, các đơn vị dự đoán, và các TU bằng cách sử dụng phép phân tích tỷ lệ biến dạng dưới dạng một phần của quy trình tối ưu hóa tỷ lệ biến dạng. Ví dụ, thành phần ước lượng chuyển động 221 có thể xác định nhiều khối tham chiếu, nhiều vector chuyển động, v.v. cho khối/khung hiện thời, và có thể chọn các khối tham chiếu, các vector chuyển động, v.v. có đặc trưng tỷ lệ biến dạng tốt nhất. Đặc trưng tỷ lệ biến dạng tốt nhất này cân bằng cả chất lượng của việc tái dựng video (ví dụ lượng tổn hao dữ liệu do nén) với hiệu quả lập mã (ví dụ kích thước của việc mã hóa cuối cùng).

Trong một số ví dụ, hệ thống lập-giải mã 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 đệm hình ảnh được giải mã 223. Ví dụ, hệ thống lập-giải mã 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 phân số khác của hình ảnh tham chiếu. Do đó, thành phần ước lượng chuyển động 221 có thể thực hiện tìm kiếm chuyển động so với các vị trí toàn điểm ảnh và các vị trí điểm ảnh phân số và xuất ra vector chuyển động có độ chính xác điểm ảnh phân số. Thành phần ước lượng chuyển động 221 tính toán vector chuyển động cho đơn vị dự đoán của khối video trong lát được lập mã liên ảnh bằng cách so sánh vị trí của đơn vị dự đoán với vị trí của khối dự đoán của hình ảnh tham chiếu. Thành phần ước lượng chuyển động 221 xuất ra vector chuyển động được tính toán dưới dạng dữ liệu chuyển động cho thành phần định dạng phần đầu và CABAC 231 để mã hóa và chuyển động đến thành phần bù chuyển động 219.

Việc bù chuyển động, được thực hiện bởi thành phần bù chuyển động 219, có thể liên quan đến việc nạp hoặc tạo ra khối dự đoán dựa vào vector chuyển động được xác

định bởi thành phần ước lượng chuyển động 221. Một lần nữa, thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219 có thể được tích hợp về mặt chức năng, theo một số ví dụ. Khi nhận vector chuyển động cho đơn vị dự đoán của khối video hiện thời, thành phần bù chuyển động 219 có thể định vị khối dự đoán đến nơi vector chuyển động chỉ vào. Sau đó, khối video dư 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 thời được lập mã, tạo ra các giá trị chênh lệch điểm ảnh. Nói chung, thành phần ước lượng chuyển động 221 thực hiện việc ước lượng chuyển động so với các thành phần độ chói, và thành phần bù chuyển động 219 sử dụng các vector chuyển động được tính toán dựa vào các thành phần độ chói cho cả các thành phần sắc độ và các thành phần độ chói. Khối dự đoán và khối dư được chuyển tiếp đến thành phần định tỷ lệ biến đổi và lượng tử hóa 213.

Tín hiệu video được phân vùng 201 còn được gửi đến thành phần ước lượng nội-hình ảnh 215 và thành phần dự đoán nội-hình ảnh 217. Cũng như thành phần ước lượng chuyển động 221 và thành phần bù chuyển động 219, thành phần ước lượng nội-hình ảnh 215 và thành phần dự đoán nội-hình ảnh 217 có thể được tích hợp ở mức độ cao, nhưng được minh họa riêng biệt cho các mục đích nhận thức. Thành phần ước lượng nội-hình ảnh 215 và thành phần dự đoán nội-hình ảnh 217 dự đoán nội ảnh khối hiện thời so với các khối trong khung hiện thời, dưới dạng thay thế cho dự đoán liên ảnh 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 nội-hình ảnh 215 xác định chế độ dự đoán nội ảnh để mã hóa khối hiện thời. Trong một số ví dụ, thành phần ước lượng nội-hình ảnh 215 lựa chọn chế độ dự đoán nội ảnh thích hợp để mã hóa khối hiện thời từ nhiều chế độ dự đoán nội ảnh được kiểm tra. Sau đó, các chế độ dự đoán nội ảnh được chuyển tiếp đến thành phần định dạng phần đầu và CABAC 231 để mã hóa.

Ví dụ, thành phần ước lượng nội-hình ảnh 215 tính toán các giá trị tỷ lệ biến dạng bằng cách sử dụng phép phân tích tỷ lệ biến dạng cho các chế độ dự đoán nội ảnh được kiểm tra khác nhau, và lựa chọn chế độ dự đoán nội ảnh có các đặc trưng tỷ lệ biến dạng tốt nhất trong số các chế độ được kiểm tra. Phép phân tích tỷ lệ biến dạng thường xác định lượng biến dạng (hoặc lỗi) giữa khối được mã hóa và khối không được mã hóa gốc được mã hóa để tạo ra khối được mã hóa, cũng như tốc độ bit (ví dụ số lượng bit) được

sử dụng để tạo ra khối được mã hóa. Thành phần ước lượng nội-hình ảnh 215 tính toán các tỷ số từ các biến dạng và các tốc độ cho các khối được mã hóa khác nhau để xác định chế độ dự đoán nội ảnh thể hiện giá trị tỷ lệ biến dạng tốt nhất cho khối này. Ngoài ra, thành phần ước lượng nội-hình ảnh 215 có thể được tạo cấu hình để lập mã các khối độ sâu của bản đồ độ sâu bằng cách sử dụng chế độ mô hình độ sâu (depth modeling mode, DMM) dựa vào sự tối ưu hóa tỷ lệ biến dạng (rate-distortion optimization, RDO).

Thành phần dự đoán nội-hình ảnh 217 có thể tạo ra khối dư từ khối dự đoán dựa vào các chế độ dự đoán nội ảnh được chọn được xác định bởi thành phần ước lượng nội-hình ảnh 215 khi được thực hiện trên bộ mã hóa hoặc đọc khối dư này từ dòng bit khi được thực hiện trên bộ giải mã. Khối dư này bao gồm chênh lệch về các giá trị giữa khối dự đoán và khối gốc, được biểu diễn dưới dạng ma trận. Sau đó, khối dư được chuyển tiếp đến thành phần định tỷ lệ biến đổi và lượng tử hóa 213. Thành phần ước lượng nội-hình ảnh 215 và thành phần dự đoán nội-hình ảnh 217 có thể hoạt động đối với cả các thành phần độ chói và sắc độ.

Thành phần định tỷ lệ biến đổi và lượng tử hóa 213 được tạo cấu hình để nén thêm khối dư. Thành phần định tỷ lệ biến đổi và lượng tử hóa 213 áp dụng phép biến đổi, như biến đổi cosin rời rạc (discrete cosine transform, DCT), biến đổi sin rời rạc (discrete sine transform, DST), hoặc phép biến đổi tương tự về khái niệm, cho khối dư, tạo ra khối video bao gồm các giá trị hệ số biến đổi dư. Các phép biến đổi sóng, phép biến đổi số nguyên, phép biến đổi dải con hoặc các loại phép biến đổi khác cũng có thể được sử dụng. Phép biến đổi này có thể biến đổi thông tin dư từ miền giá trị điểm ảnh sang miền biến đổi, như miền tần số. Thành phần định tỷ lệ biến đổi và lượng tử hóa 213 cũng được tạo cấu hình để định tỷ lệ thông tin dư được biến đổi, ví dụ dựa vào tần số. Việc định tỷ lệ này bao gồm việc áp dụng hệ số định tỷ lệ cho thông tin dư sao cho thông tin tần số khác được lượng tử hóa ở các độ chi tiết khác nhau, có thể ảnh hưởng đến chất lượng thị giác cuối cùng của video được tái dựng. Thành phần định tỷ lệ biến đổi và lượng tử hóa 213 được tạo cấu hình để lượng tử hóa các hệ số biến đổi để giảm thêm tốc độ bit. Quy trình lượng tử hóa có thể giảm độ sâu bit liên quan 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 số ví dụ, sau đó thành phần định tỷ lệ biến đổi và lượng tử hóa 213 có thể 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 này được chuyển tiếp đến thành phần định dạng phần đầu và CABAC 231 để được mã hóa trong dòng bit.

Thành phần định tỷ lệ và biến đổi ngược 229 áp dụng hoạt động ngược với thành phần định tỷ lệ biến đổi và lượng tử hóa 213 để hỗ trợ việc ước lượng chuyển động. Thành phần định tỷ lệ và biến đổi ngược 229 áp dụng việc định tỷ lệ, biến đổi, và/hoặc lượng tử hóa ngược để tái dựng khối dư trong miền điểm ảnh, ví dụ, để sử dụng sau đó làm khối tham chiếu mà có thể trở thành khối dự đoán cho khối hiện thời 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 bổ sung khối dư trở lại khối dự đoán tương ứng để sử dụng trong quá trình ước lượng chuyển động của khối/khung sau. Các bộ lọc được áp dụng cho các khối tham chiếu được tái dựng để làm giảm thiểu các nhiễu được tạo ra trong quá trình định tỷ lệ, lượng tử hóa, và biến đổi. Mặt khác, các nhiễu này có thể khiến phép dự đoán không chính xác (và tạo ra các nhiễu bổ sung) khi các khối sau đó được dự đoán.

Thành phần phân tích điều khiển bộ lọc 227 và thành phần lọc trong vòng lặp 225 áp dụng các bộ lọc cho các khối dư và/hoặc cho các khối ảnh được tái dựng. Ví dụ, khối dư được biến đổi từ thành phần định tỷ lệ và biến đổi ngược 229 có thể được kết hợp với khối dự đoán tương ứng từ thành phần dự đoán nội-hình ảnh 217 và/hoặc thành phần bù chuyển động 219 để tái dựng khối ảnh gốc. Sau đó, các bộ lọc có thể được áp dụng cho khối ảnh được tái dựng. Trong một số ví dụ, các bộ lọc có thể được áp dụng thay cho các khối dư. Cũng như các thành phần khác trên FIG.2, thành phần phân tích điều khiển bộ lọc 227 và thành phần lọc trong vòng lặp 225 được tích hợp ở mức độ cao và có thể được thực hiện cùng nhau, nhưng được thể hiện riêng biệt cho các mục đích nhận thức. Các bộ lọc được áp dụng cho các khối tham chiếu được tái dựng được áp dụng cho các vùng theo không gian cụ thể và bao gồm nhiều tham số để điều chỉnh cách các bộ lọc được áp dụng. Thành phần phân tích điều khiển bộ lọc 227 phân tích các khối tham chiếu được tái dựng để xác định nơi mà các bộ lọc cần được áp dụng và thiết lập các tham số tương ứng. Dữ liệu này được chuyển tiếp đến thành phần định dạng phần đầu và CABAC 231 dưới dạng dữ liệu điều khiển bộ lọc để mã hóa. Thành phần lọc trong vòng lặp 225 áp dụng các bộ lọc này dựa vào dữ liệu điều khiển bộ lọc. Các bộ lọc có thể bao gồm bộ lọc tách khối, bộ lọc khử nhiễu, bộ lọc SAO, và bộ lọc vòng lặp thích

ứng. Các bộ lọc này có thể được áp dụng trong miền theo không gian/điểm ảnh (ví dụ, trên khối điểm ảnh được tái dựng) hoặc trong miền tần số, tùy thuộc vào ví dụ.

Khi hoạt động như bộ mã hóa, khối ảnh tái dựng được lọc, khối dư, và/hoặc khối dự đoán được lưu trữ trong thành phần đệm hình ảnh được giải mã 223 để sử dụng sau đó trong quá trình ước lượng chuyển động như được trình bày ở trên. Khi hoạt động như bộ giải mã, thành phần đệm hình ảnh được giải mã 223 lưu trữ và chuyển tiếp các khối được tái dựng và được lọc về phía màn hình dưới dạng một phần của tín hiệu video đầu ra. Thành phần đệm hình ảnh được giải mã 223 có thể là thiết bị nhớ bất kỳ có khả năng lưu trữ các khối dự đoán, các khối dư, và/hoặc các khối ảnh được tái dựng.

Thành phần định dạng phần đầu và CABAC 231 nhận dữ liệu từ các thành phần khác nhau của hệ thống lập-giải mã 200 và mã hóa dữ liệu này thành dòng bit được lập mã để truyền về phía bộ giải mã. Cụ thể là, thành phần định dạng phần đầu và CABAC 231 tạo ra các phần đầu khác nhau để mã hóa dữ liệu điều khiển, như dữ liệu điều khiển chung và dữ liệu điều khiển bộ lọc. Ngoài ra, dữ liệu dự đoán, bao gồm dự đoán nội ảnh và dữ liệu chuyển động, cũng như dữ liệu dư dưới 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 các thông tin mong muốn bởi bộ giải mã để tái dựng tín hiệu video được phân vùng gốc 201. Thông tin này cũng có thể bao gồm các bảng chỉ số chế độ dự đoán nội ảnh (cũng được gọi là các bảng ánh xạ từ mã), các định nghĩa của các ngữ cảnh mã hóa cho các khối khác nhau, các chỉ báo của các chế độ dự đoán nội ảnh có khả năng xảy ra nhất, chỉ báo về thông tin phân vùng, v.v.. Dữ liệu này có thể được mã hóa bằng cách sử dụng lập mã entropi. Ví dụ, thông tin có thể được mã hóa bằng cách sử dụng lập mã độ dài biến đổi thích ứng ngữ cảnh (context adaptive variable length coding, CAVLC), CABAC, lập mã số học nhị phân thích ứng ngữ cảnh dựa vào cú pháp (syntax-based context-adaptive binary arithmetic coding, SBAC), lập mã entropi phân vùng khoảng thời gian có thể xảy ra (probability interval partitioning entropy, PIPE), hoặc kỹ thuật lập mã entropi khác. Sau khi lập mã entropi, dòng bit được lập mã có thể được truyền đến thiết bị khác (ví dụ, bộ giải mã video) hoặc được lưu để truyền dẫn hoặc truy xuất sau này.

FIG.3 là sơ đồ khối minh họa bộ mã hóa video 300 làm ví dụ. Bộ mã hóa video 300 có thể được sử dụng để thực hiện các chức năng mã hóa của hệ thống lập-giải mã 200 và/hoặc thực hiện các bước 101, 103, 105, 107, và/hoặc 109 của phương pháp hoạt

động 100. Bộ mã hóa 300 phân vùng tín hiệu video đầu vào, tạo ra tín hiệu video được phân vùng 301, gần như tương tự tín hiệu video được phân vùng 201. Sau đó, tín hiệu video được phân vùng 301 được nén và được mã hóa thành dòng bit bởi các thành phần của bộ mã hóa 300.

Cụ thể là, tín hiệu video được phân vùng 301 được chuyển tiếp đến thành phần dự đoán nội-hình ảnh 317 để dự đoán nội ảnh. Thành phần dự đoán nội-hình ảnh 317 có thể gần như tương tự thành phần ước lượng nội-hình ảnh 215 và thành phần dự đoán nội-hình ảnh 217. Tín hiệu video được phân vùng 301 còn được chuyển tiếp đến thành phần bù chuyển động 321 để dự đoán liên ảnh dựa vào các khối tham chiếu trong thành phần đệm hình ảnh được giải mã 323. Thành phần bù chuyển động 321 có thể gần như tương tự 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 nội-hình ảnh 317 và thành phần bù chuyển động 321 được chuyển tiếp đến thành phần biến đổi và lượng tử hóa 313 để biến đổi và lượng tử hóa của các khối dư. Thành phần biến đổi và lượng tử hóa 313 có thể gần như tương tự thành phần định tỷ lệ biến đổi và lượng tử hóa 213. Các khối dư được biến đổi và được lượng tử hóa và các khối dự đoán tương ứng (cùng với dữ liệu điều khiển có liên quan) được chuyển tiếp đến thành phần lập mã entropi 331 để lập mã thành dòng bit. Thành phần lập mã entropi 331 có thể gần như tương tự thành phần định dạng phần đầu và CABAC 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 và lượng tử hóa ngược 329 để tái dựng thành các khối tham chiếu để sử dụng bởi thành phần bù chuyển động 321. Thành phần biến đổi và lượng tử hóa ngược 329 có thể gần như tương tự thành phần định tỷ lệ và biến đổi ngược 229. Các bộ lọc trong vòng lặp trong thành phần lọc trong vòng lặp 325 cũng được áp dụng cho các khối dư và/hoặc các khối tham chiếu được tái dựng, tùy thuộc vào ví dụ. Thành phần lọc trong vòng lặp 325 có thể gần như tương tự thành phần phân tích điều khiển bộ lọc 227 và thành phần lọc trong vòng lặp 225. Thành phần lọc trong vòng lặp 325 có thể bao gồm nhiều bộ lọc như được trình bày đối với thành phần lọc trong vòng lặp 225. Sau đó, các khối được lọc được lưu trữ trong thành phần đệm hình ảnh được giải mã 323 để sử dụng

làm các khối tham chiếu bởi thành phần bù chuyển động 321. Thành phần đệm hình ảnh được giải mã 323 có thể gần như tương tự thành phần đệm hình ảnh được giải mã 223.

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

Các dòng bit được tiếp nhận bởi thành phần giải mã entropi 433. Thành phần giải mã entropi 433 được tạo cấu hình để thực hiện sơ đồ giải mã entropi, như lập mã CAVLC, CABAC, SBAC, PIPE, hoặc các kỹ thuật lập mã entropi khác. Ví dụ, thành phần giải mã entropi 433 có thể sử dụng thông tin phần đầu để tạo ngữ cảnh để diễn giải dữ liệu bổ sung được mã hóa dưới dạng 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 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 và lượng tử hóa ngược 429 để tái dựng thành các khối dư. Thành phần biến đổi và lượng tử hóa ngược 429 có thể tương tự thành phần biến đổi và lượng tử hóa ngược 329.

Các khối dư được tái dựng và/hoặc các khối dự đoán được chuyển tiếp đến thành phần dự đoán nội-hình ảnh 417 để tái dựng thành các khối ảnh dựa vào các hoạt động dự đoán nội ảnh. Thành phần dự đoán nội-hình ảnh 417 có thể tương tự thành phần ước lượng nội-hình ảnh 215 và thành phần dự đoán nội-hình ảnh 217. Cụ thể là, thành phần dự đoán nội-hình ảnh 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 tái dựng các khối ảnh được dự đoán nội ảnh. Các khối ảnh được dự đoán nội ảnh được tái dựng và/hoặc các khối dư và dữ liệu dự đoán liên ảnh tương ứng được chuyển tiếp đến thành phần đệm hình ảnh được giải mã 423 thông qua thành phần lọc trong vòng lặp 425, có thể gần như tương tự thành phần đệm hình ảnh được giải mã 223 và thành phần lọc trong vòng lặp 225, một cách tương ứng. Thành phần lọc trong vòng lặp 425 lọc các khối ảnh được tái dựng, các khối dư và/hoặc các khối dự đoán, và thông tin này được lưu trữ trong thành phần đệm hình ảnh được giải mã 423. Các khối ảnh được tái dựng từ thành phần đệm hình ảnh được giải

mã 423 được chuyển tiếp đến thành phần bù chuyển động 421 để dự đoán liên ảnh. Thành phần bù chuyển động 421 có thể gần như tương tự thành phần ước lượng chuyển động 221 và/hoặc thành phần bù chuyển động 219. Cụ thể là, thành phần bù chuyển động 421 sử dụng các vectơ chuyển động từ khối tham chiếu để tạo ra khối dự đoán và áp dụng khối dư cho kết quả này để tái dựng khối ảnh. Các khối được tái dựng thu được cũng có thể được chuyển tiếp thông qua thành phần lọc trong vòng lặp 425 đến thành phần đệm hình ảnh được giải mã 423. Thành phần đệm hình ảnh được giải mã 423 tiếp tục lưu trữ các khối ảnh được tái dựng bổ sung, có thể được tái dựng thành các khung thông qua thông tin phân vùng. Các khung này cũng có thể được đặt theo trình tự. Trình tự này được xuất ra về phía màn hình dưới dạng tín hiệu video đầu ra được tái dựng.

FIG.5 là sơ đồ giản lược minh họa HRD 500 làm ví dụ. HRD 500 có thể được sử dụng trong bộ mã hóa, như hệ thống lập-giải mã 200 và/hoặc bộ mã hóa 300. HRD 500 có thể kiểm tra dòng bit được tạo ra ở bước 109 của phương pháp 100 trước khi dòng bit được chuyển tiếp đến bộ giải mã, như bộ giải mã 400. Trong một số ví dụ, dòng bit có thể được chuyển tiếp liên tục qua HRD 500 dưới dạng dòng bit được mã hóa. Trong trường hợp một phần của dòng bit không phù hợp với các ràng buộc có liên quan, HRD 500 có thể biểu thị lỗi này cho bộ mã hóa để khiến bộ mã hóa mã hóa lại đoạn tương ứng của dòng bit bằng các cơ chế khác nhau.

HRD 500 bao gồm bộ lập lịch dòng giả định (hypothetical stream scheduler, HSS) 541. HSS 541 là thành phần được tạo cấu hình để thực hiện cơ chế phân phối giả định. Cơ chế phân phối giả định được sử dụng để kiểm tra sự tương thích của dòng bit hoặc bộ giải mã đối với việc định thời và dòng dữ liệu của dòng bit 551 được nhập vào HRD 500. Ví dụ, HSS 541 có thể nhận dòng bit 551 được xuất ra từ bộ mã hóa và quản lý quy trình kiểm tra sự tương thích trên dòng bit 551. Trong một ví dụ cụ thể, HSS 541 có thể điều khiển tốc độ mà các hình ảnh được lập mã di chuyển qua HRD 500 và xác nhận rằng dòng bit 551 không chứa dữ liệu không thích hợp.

HSS 541 có thể chuyển tiếp dòng bit 551 đến CPB 543 ở tốc độ định trước. HRD 500 có thể quản lý dữ liệu trong các đơn vị giải mã (decoding unit, DU) 553. DU 553 là đơn vị truy cập (access unit, AU) hoặc tập hợp con của AU và các đơn vị lớp trừu tượng hóa mạng loại lớp lập mã video (video coding layer (VCL) network abstraction layer (NAL) unit) không được liên kết. Cụ thể là, AU chứa một hoặc nhiều hình ảnh được liên

kết với thời gian đầu ra. Ví dụ, AU có thể chứa một hình ảnh trong dòng bit một lớp, và có thể chứa hình ảnh cho mỗi lớp trong dòng bit đa lớp. Mỗi hình ảnh của AU có thể được chia thành các lát đều có trong đơn vị VCL NAL tương ứng. Do đó, DU 553 có thể chứa một hoặc nhiều hình ảnh, một hoặc nhiều lát của hình ảnh, hoặc các tổ hợp của chúng. Ngoài ra, các tham số được sử dụng để giải mã AU, các hình ảnh, và/hoặc các lát có thể có trong các đơn vị non-VCL NAL. Do đó, DU 553 chứa các đơn vị non-VCL NAL chứa dữ liệu cần hỗ trợ việc giải mã các đơn vị VCL NAL trong DU 553. CPB 543 là bộ đệm vào trước ra trước trong HRD 500. CPB 543 chứa các DU 553 bao gồm dữ liệu video theo thứ tự giải mã. CPB 543 lưu trữ dữ liệu video để sử dụng trong quá trình xác nhận sự tương thích dòng bit.

CPB 543 chuyển tiếp các DU 553 đến thành phần xử lý giải mã 545. Thành phần xử lý giải mã 545 là thành phần thích hợp với chuẩn VVC. Ví dụ, thành phần xử lý giải mã 545 có thể mô phỏng bộ giải mã 400 được sử dụng bởi người dùng đầu cuối. Thành phần xử lý giải mã 545 giải mã các DU 553 ở tốc độ có thể đạt được bởi bộ giải mã người dùng đầu cuối làm ví dụ. Nếu thành phần xử lý giải mã 545 không thể giải mã các DU 553 đủ nhanh để ngăn chặn được sự tràn CPB 543, thì dòng bit 551 không thích hợp với chuẩn này và cần được mã hóa lại.

Thành phần xử lý giải mã 545 giải mã các DU 553, tạo ra các DU được giải mã 555. DU được giải mã 555 chứa hình ảnh được giải mã. Các DU được giải mã 555 được chuyển tiếp đến DPB 547. DPB 547 có thể gần như tương tự thành phần đệm hình ảnh được giải mã 223, 323, và/hoặc 423. Để hỗ trợ dự đoán liên ảnh, các hình ảnh được đánh dấu để sử dụng làm các hình ảnh tham chiếu 556 được thu nhận từ các DU được giải mã 555 được đưa trở lại thành phần xử lý giải mã 545 để hỗ trợ việc giải mã khác. DPB 547 xuất ra chuỗi video được giải mã dưới dạng chuỗi hình ảnh 557. Các hình ảnh 557 là các hình ảnh được tái dựng thường là các hình ảnh gương được mã hóa thành dòng bit 551 bởi bộ mã hóa.

Các hình ảnh 557 được chuyển tiếp đến thành phần cắt xén đầu ra 549. Thành phần cắt xén đầu ra 549 được tạo cấu hình để áp dụng cửa sổ cắt xén tương thích cho các hình ảnh 557. Việc này tạo ra các hình ảnh được cắt xén đầu ra 559. Hình ảnh được cắt xén đầu ra 559 là hình ảnh được tái dựng hoàn toàn. Do đó, hình ảnh được cắt xén đầu ra 559 mô phỏng lại những gì người dùng đầu cuối sẽ nhìn thấy khi giải mã dòng

bit 551. Do đó, bộ mã hóa có thể xem xét lại các hình ảnh được cắt xén đầu ra 559 để bảo đảm việc mã hóa là thỏa đáng.

HRD 500 được kích hoạt dựa vào các tham số HRD trong dòng bit 551. Ví dụ, HRD 500 có thể đọc các tham số HRD từ VPS, SPS, và/hoặc các thông điệp SEI. Sau đó, HRD 500 có thể thực hiện các hoạt động kiểm tra sự tương thích trên dòng bit 551 dựa vào thông tin trong các tham số HRD này. Trong một ví dụ cụ thể, HRD 500 có thể xác định một hoặc nhiều lịch phân phối CPB từ các tham số HRD. Lịch phân phối này xác định việc định thời cho việc phân phối dữ liệu video đến và/hoặc từ vị trí bộ nhớ, như CPB và/hoặc DPB. Do đó, lịch phân phối CPB xác định việc định thời cho việc phân phối của các AU, các DU 553, và/hoặc các hình ảnh, đến/từ CPB 543. Cần lưu ý rằng HRD 500 có thể sử dụng các lịch phân phối DPB cho DPB 547 tương tự các lịch phân phối CPB.

Video có thể được lập mã thành các lớp và/hoặc các OLS khác nhau để sử dụng bởi các bộ giải mã có các cấp độ khả năng phần cứng khác nhau cũng như cho các điều kiện mạng khác nhau. Các lịch phân phối CPB được chọn để phản ánh các vấn đề này. Do đó, các dòng bit con lớp cao hơn được chỉ định cho các điều kiện phần cứng và mạng tối ưu và do đó các lớp cao hơn có thể nhận một hoặc nhiều lịch phân phối CPB sử dụng lượng bộ nhớ lớn trong CPB 543 và độ trễ ngắn cho sự truyền dẫn của các DU 553 về phía DPB 547. Tương tự, các dòng bit con lớp dưới được chỉ định cho các khả năng phần cứng bộ giải mã giới hạn và/hoặc các điều kiện mạng kém. Do đó, các lớp dưới có thể nhận một hoặc nhiều lịch phân phối CPB sử dụng lượng bộ nhớ nhỏ trong CPB 543 và các độ trễ dài hơn cho sự truyền dẫn của các DU 553 về phía DPB 547. Sau đó, các OLS, các lớp, các lớp con hoặc các tổ hợp của chúng có thể được kiểm tra theo lịch phân phối tương ứng để bảo đảm rằng dòng bit con thu được có thể được giải mã đúng dưới các điều kiện được mong muốn cho dòng bit con. Do đó, các tham số HRD trong dòng bit 551 có thể biểu thị các lịch phân phối CPB cũng như bao gồm dữ liệu đủ để cho phép HRD 500 xác định các lịch phân phối CPB và phối hợp các lịch phân phối CPB với các OLS, các lớp và/hoặc các lớp con tương ứng.

FIG.6 là sơ đồ giản lược minh họa chuỗi video đa lớp 600 làm ví dụ. Chuỗi video đa lớp 600 có thể được mã hóa bởi bộ mã hóa, như hệ thống lập-giải mã 200 và/hoặc bộ mã hóa 300 và được giải mã bởi bộ giải mã, như hệ thống lập-giải mã 200 và/hoặc bộ

giải mã 400, ví dụ theo phương pháp 100. Ngoài ra, chuỗi video đa lớp 600 cũng có thể được kiểm tra về sự tương thích với chuẩn bởi HRD, như HRD 500. Chuỗi video đa lớp 600 được bao gồm để mô tả ứng dụng làm ví dụ cho các lớp trong chuỗi video được lập mã. Chuỗi video đa lớp 600 là chuỗi video bất kỳ sử dụng nhiều lớp, như lớp N 631 và lớp N+1 632.

Trong ví dụ, chuỗi video đa lớp 600 có thể sử dụng phép dự đoán liên-lớp 621. Phép dự đoán liên-lớp 621 được áp dụng giữa các hình ảnh 611, 612, 613, và 614 và các hình ảnh 615, 616, 617, và 618 trong các lớp khác nhau. Trong ví dụ được thể hiện, các hình ảnh 611, 612, 613, và 614 là một phần của lớp N+1 632 và các hình ảnh 615, 616, 617, và 618 là một phần của lớp N 631. Một lớp, như lớp N 631 và/hoặc lớp N+1 632, là nhóm gồm các hình ảnh đều được liên kết với một giá trị tương tự của một đặc trưng, như kích thước, chất lượng, độ phân giải, tỷ số tín hiệu trên nhiễu, dung lượng, v.v. tương tự. Một lớp có thể được xác định chính thức dưới dạng tập hợp các đơn vị VCL NAL chia sẻ cùng ID lớp và các đơn vị non-VCL NAL có liên quan. Đơn vị VCL NAL là đơn vị NAL được lập mã để chứa dữ liệu video, như lát được lập mã của hình ảnh. Đơn vị non-VCL NAL là đơn vị NAL chứa dữ liệu không phải video như cú pháp và/hoặc các tham số hỗ trợ việc giải mã dữ liệu video, hiệu suất của phép kiểm tra sự tương thích, hoặc các hoạt động khác.

Trong ví dụ được thể hiện, lớp N+1 632 được liên kết với kích thước hình ảnh lớn hơn so với lớp N 631. Do đó, các hình ảnh 611, 612, 613, và 614 trong lớp N+1 632 có kích thước hình ảnh lớn hơn (ví dụ, chiều cao và chiều rộng lớn hơn và do đó có nhiều mẫu hơn) so với các hình ảnh 615, 616, 617, và 618 trong lớp N 631 trong ví dụ này. Tuy nhiên, các hình ảnh này có thể được tách rời giữa lớp N+1 632 và lớp N 631 bởi các đặc trưng khác. Mặc dù chỉ có hai lớp, lớp N+1 632 và lớp N 631, như được thể hiện trên hình vẽ, nhưng tập hợp các hình ảnh có thể được tách rời thành số lượng lớp bất kỳ dựa vào các đặc trưng có liên quan. Lớp N+1 632 và lớp N 631 cũng có thể được biểu thị bởi ID lớp. ID lớp là một mục của dữ liệu được liên kết với hình ảnh và biểu thị hình ảnh này là một phần của lớp được biểu thị. Do đó, mỗi hình ảnh 611-618 có thể được liên kết với ID lớp tương ứng để biểu thị lớp N+1 632 hoặc lớp N 631 bao gồm hình ảnh tương ứng. Ví dụ, ID lớp có thể bao gồm mã định danh lớp phần đầu đơn vị NAL (nuh_layer_id), là phần tử cú pháp xác định mã định danh của lớp bao gồm đơn vị

NAL (ví dụ, bao gồm các lát và/hoặc các tham số của các hình ảnh trong lớp). Lớp liên quan đến chất lượng thấp hơn/kích thước hình ảnh nhỏ hơn/kích thước dòng bit nhỏ hơn, như lớp N 631, thường được gán ID lớp dưới và được gọi là lớp dưới. Ngoài ra, lớp liên quan đến chất lượng cao hơn/kích thước hình ảnh lớn hơn/kích thước dòng bit lớn hơn, như lớp N+1 632, thường được gán ID lớp cao hơn và được gọi là lớp cao hơn.

Các hình ảnh 611-618 trong các lớp khác nhau 631-632 được tạo cấu hình để được hiển thị theo cách luân phiên. Là một ví dụ cụ thể, bộ giải mã có thể giải mã và hiển thị hình ảnh 615 ở thời gian hiển thị hiện thời nếu mong muốn hình ảnh nhỏ hơn hoặc bộ giải mã có thể giải mã và hiển thị hình ảnh 611 ở thời gian hiển thị hiện thời nếu mong muốn hình ảnh lớn hơn. Do đó, các hình ảnh 611-614 ở lớp cao hơn N+1 632 về cơ bản chứa cùng dữ liệu hình ảnh như các hình ảnh tương ứng 615-618 ở lớp dưới N 631 (bất kể sự chênh lệch trong kích thước hình ảnh). Cụ thể là, hình ảnh 611 về cơ bản chứa cùng dữ liệu hình ảnh như hình ảnh 615, hình ảnh 612 về cơ bản chứa cùng dữ liệu hình ảnh như hình ảnh 616 v.v..

Các hình ảnh 611-618 có thể được lập mã bằng cách tham chiếu đến các hình ảnh khác 611-618 trong cùng lớp N 631 hoặc N+1 632. Việc lập mã hình ảnh dựa vào hình ảnh khác trong cùng lớp dẫn đến phép dự đoán liên ảnh 623. Phép dự đoán liên ảnh 623 được thể hiện bằng các mũi tên nét liền. Ví dụ, hình ảnh 613 có thể được lập mã bằng cách sử dụng phép dự đoán liên ảnh 623 sử dụng một hoặc hai trong số các hình ảnh 611, 612, và/hoặc 614 trong lớp N+1 632 làm tham chiếu, trong đó một hình ảnh được tham chiếu cho dự đoán liên ảnh một chiều và/hoặc hai hình ảnh được tham chiếu cho dự đoán liên ảnh hai chiều. Ngoài ra, hình ảnh 617 có thể được lập mã bằng cách sử dụng phép dự đoán liên ảnh 623 sử dụng một hoặc hai hình ảnh 615, 616, và/hoặc 618 trong lớp N 631 làm tham chiếu, trong đó một hình ảnh được tham chiếu cho dự đoán liên ảnh một chiều và/hoặc hai hình ảnh được tham chiếu cho dự đoán liên ảnh hai chiều. Khi hình ảnh được sử dụng làm tham chiếu cho hình ảnh khác trong cùng lớp khi thực hiện dự đoán liên ảnh 623, hình ảnh này có thể được gọi là hình ảnh tham chiếu. Ví dụ, hình ảnh 612 có thể là hình ảnh tham chiếu được sử dụng để lập mã hình ảnh 613 theo dự đoán liên ảnh 623. Dự đoán liên ảnh 623 cũng có thể được gọi là phép dự đoán nội-lớp trong ngữ cảnh đa lớp. Do đó, dự đoán liên ảnh 623 là cơ chế lập mã các mẫu của hình ảnh hiện thời bằng cách tham chiếu đến các mẫu được biểu thị trong hình ảnh tham

chiều khác với hình ảnh hiện thời trong đó hình ảnh tham chiếu và hình ảnh hiện thời nằm trong cùng lớp.

Các hình ảnh 611-618 cũng có thể được lập mã bằng cách tham chiếu đến các hình ảnh khác 611-618 trong các lớp khác nhau. Quy trình này đã được biết là phép dự đoán liên-lớp 621, và được thể hiện bằng các mũi tên nét đứt. Phép dự đoán liên-lớp 621 là cơ chế lập mã các mẫu của hình ảnh hiện thời bằng cách tham chiếu đến các mẫu được biểu thị trong hình ảnh tham chiếu là nơi hình ảnh hiện thời và hình ảnh tham chiếu nằm trong các lớp khác nhau và do đó có các ID lớp khác nhau. Ví dụ, hình ảnh trong lớp dưới N 631 có thể được sử dụng làm hình ảnh tham chiếu để lập mã hình ảnh tương ứng ở lớp cao hơn N+1 632. Trong một ví dụ cụ thể, hình ảnh 611 có thể được lập mã bằng cách tham chiếu đến hình ảnh 615 theo phép dự đoán liên-lớp 621. Trong trường hợp này, hình ảnh 615 được sử dụng làm hình ảnh tham chiếu liên-lớp. Hình ảnh tham chiếu liên-lớp là hình ảnh tham chiếu được sử dụng cho phép dự đoán liên-lớp 621. Trong hầu hết các trường hợp, phép dự đoán liên-lớp 621 được ràng buộc sao cho hình ảnh hiện thời, như hình ảnh 611, có thể chỉ sử dụng (các) hình ảnh tham chiếu liên-lớp có trong cùng AU 627 và ở lớp dưới, như hình ảnh 615. Khi nhiều lớp (ví dụ, nhiều hơn hai) khả dụng, phép dự đoán liên-lớp 621 có thể mã hóa/giải mã hình ảnh hiện thời dựa vào nhiều hình ảnh tham chiếu liên-lớp ở các cấp thấp hơn so với hình ảnh hiện thời.

Bộ mã hóa video có thể sử dụng chuỗi video đa lớp 600 để mã hóa các hình ảnh 611-618 thông qua nhiều tổ hợp và/hoặc sự hoán vị khác nhau của dự đoán liên ảnh 623 và phép dự đoán liên-lớp 621. Ví dụ, hình ảnh 615 có thể được lập mã theo dự đoán nội ảnh. Sau đó, các hình ảnh 616-618 có thể được lập mã theo dự đoán liên ảnh 623 bằng cách sử dụng hình ảnh 615 làm hình ảnh tham chiếu. Ngoài ra, hình ảnh 611 có thể được lập mã theo phép dự đoán liên-lớp 621 bằng cách sử dụng hình ảnh 615 làm hình ảnh tham chiếu liên-lớp. Sau đó, các hình ảnh 612-614 có thể được lập mã theo dự đoán liên ảnh 623 bằng cách sử dụng hình ảnh 611 làm hình ảnh tham chiếu. Do đó, hình ảnh tham chiếu có thể dành cho cả hình ảnh tham chiếu một lớp lẫn hình ảnh tham chiếu liên-lớp cho các cơ chế lập mã khác nhau. Bằng cách lập mã các hình ảnh lớp cao hơn N+1 632 dựa vào các hình ảnh lớp dưới N 631, lớp cao hơn N+1 632 có thể tránh sử dụng dự đoán nội ảnh, có hiệu quả lập mã thấp hơn nhiều so với phép dự đoán liên ảnh 623 và phép dự đoán liên-lớp 621. Do đó, hiệu quả lập mã thấp của phép dự đoán nội ảnh có

thể bị giới hạn ở các hình ảnh có chất lượng nhỏ nhất/thấp nhất, và do đó bị giới hạn ở việc lập mã lượng dữ liệu video nhỏ nhất. Các hình ảnh này được sử dụng làm các hình ảnh tham chiếu và/hoặc các hình ảnh tham chiếu liên-lớp có thể được biểu thị trong các đầu vào của (các) danh sách hình ảnh tham chiếu chứa trong cấu trúc danh sách hình ảnh tham chiếu.

Các hình ảnh 611-618 cũng có thể có trong các đơn vị truy cập (các AU) 627. AU 627 là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian đầu ra trong quá trình giải mã. Do đó, các hình ảnh được lập mã trong cùng AU 627 được lập lịch cho đầu ra từ DPB ở bộ giải mã đồng thời. Ví dụ, các hình ảnh 614 và 618 trong cùng AU 627. Các hình ảnh 613 và 617 nằm trong một AU 627 khác từ các hình ảnh 614 và 618. Các hình ảnh 614 và 618 trong cùng AU 627 có thể được hiển thị theo cách luân phiên. Ví dụ, hình ảnh 618 có thể được hiển thị khi mong muốn kích thước hình ảnh nhỏ và hình ảnh 614 có thể được hiển thị khi mong muốn kích thước hình ảnh lớn. Khi mong muốn kích thước hình ảnh lớn, hình ảnh 614 được xuất ra và hình ảnh 618 chỉ được sử dụng cho phép dự đoán liên lớp 621. Trong trường hợp này, hình ảnh 618 được loại bỏ mà không xuất ra mỗi khi phép dự đoán liên lớp 621 hoàn thành.

AU 627 cũng có thể được chia thành một hoặc nhiều đơn vị hình ảnh (PU) 628. PU 628 là tập hợp con của AU 627 chứa một hình ảnh được lập mã. PU 628 có thể được xác định chính thức dưới dạng tập hợp các đơn vị NAL được liên kết với nhau theo quy tắc phân loại xác định, là liên tiếp theo thứ tự giải mã, và chứa chính xác một hình ảnh được lập mã. Cần lưu ý rằng PU 628 có thể được gọi là đơn vị giải mã (DU) khi được mô tả đối với HRD và/hoặc các phép kiểm tra sự tương thích có liên quan.

Ngoài ra, cần lưu ý rằng các hình ảnh 611-618, và do đó các AU 627 và các PU 628, đều được liên kết với mã định danh theo thời gian (TemporalId) 629. TemporalId 629 là mã định danh được suy ra biểu thị vị trí tương đối của đơn vị NAL trong chuỗi video. Các hình ảnh và/hoặc các PU 628 trong cùng AU 627 được liên kết với cùng giá trị của TemporalId 629. Ví dụ, AU thứ nhất 627 theo trình tự có thể bao gồm TemporalId 629 bằng zero, với các AU 627 sau đó bao gồm các TemporalId 629 tăng liên tiếp. Các đơn vị non-VCL NAL cũng có thể được liên kết với các TemporalId 629. Ví dụ, một tập hợp tham số có thể có trong AU 627 và có thể được liên kết với một hoặc nhiều hình

ảnh trong AU 627. Trong trường hợp này, TemporalId 629 của tập hợp tham số có thể nhỏ hơn hoặc bằng TemporalId 629 của AU 627.

FIG.7 là sơ đồ giản lược minh họa dòng bit 700 làm ví dụ. Ví dụ, dòng bit 700 có thể được tạo ra bởi hệ thống lập-giải mã 200 và/hoặc bộ mã hóa 300 để giải mã bởi hệ thống lập-giải mã 200 và/hoặc bộ giải mã 400 theo phương pháp 100. Ngoài ra, dòng bit 700 có thể bao gồm chuỗi video đa lớp 600. Ngoài ra, dòng bit 700 có thể bao gồm các tham số khác nhau để điều khiển hoạt động của HRD, như HRD 500. Dựa vào các tham số này, HRD có thể kiểm tra dòng bit 700 về sự tương thích với các chuẩn trước khi truyền về phía bộ giải mã để giải mã.

Dòng bit 700 bao gồm VPS 711, một hoặc nhiều SPS 713, các tập hợp tham số hình ảnh (PPS) 715, các tập hợp tham số thích ứng (APS) 716, các phần đầu hình ảnh 718, các phần đầu lát 717, dữ liệu hình ảnh 720, và các thông điệp SEI 719. VPS 711 chứa dữ liệu có liên quan đến toàn bộ dòng bit 700. Ví dụ, VPS 711 có thể chứa dữ liệu liên quan đến các OLS, các lớp, và/hoặc các lớp con được sử dụng trong dòng bit 700. SPS 713 chứa dữ liệu trình tự chung cho tất cả các hình ảnh trong chuỗi video được lập mã chứa trong dòng bit 700. Ví dụ, mỗi lớp có thể chứa một hoặc nhiều chuỗi video được lập mã, và mỗi chuỗi video được lập mã có thể tham chiếu đến SPS 713 cho các tham số tương ứng. Các tham số trong SPS 713 có thể bao gồm kích thước hình ảnh, độ sâu bit, các tham số công cụ lập mã, các giới hạn tốc độ bit v.v.. Cần lưu ý rằng, mặc dù mỗi trình tự đề cập đến SPS 713, nhưng một SPS 713 có thể chứa dữ liệu cho nhiều trình tự trong một số ví dụ. PPS 715 chứa các tham số áp dụng cho hình ảnh nguyên vẹn. Do đó, mỗi hình ảnh trong chuỗi video có thể đề cập đến PPS 715. Cần lưu ý rằng, mặc dù mỗi hình ảnh đề cập đến PPS 715, nhưng một PPS 715 có thể chứa dữ liệu cho nhiều hình ảnh trong một số ví dụ. Ví dụ, nhiều hình ảnh tương tự có thể được lập mã theo các tham số tương tự. Trong trường hợp này, một PPS 715 có thể chứa dữ liệu cho các hình ảnh tương tự này. PPS 715 có thể biểu thị các công cụ lập mã khả dụng cho các lát trong các hình ảnh tương ứng, các tham số lượng tử hóa, các độ lệch, v.v..

APS 716 là cấu trúc cú pháp chứa các phần tử cú pháp/các tham số áp dụng cho một hoặc nhiều lát 727 trong một hoặc nhiều hình ảnh 725. Sự tương quan này có thể được xác định dựa vào các phần tử cú pháp được tìm thấy trong các phần đầu lát 717 được liên kết với các lát 727. Ví dụ, APS 716 có thể áp dụng cho ít nhất một, nhưng

không phải tất cả, các lát 727 trong hình ảnh thứ nhất 721, cho ít nhất một, nhưng không phải tất cả, các lát 727 trong hình ảnh thứ hai 725, v.v.. APS 716 có thể được chia thành nhiều kiểu dựa vào các tham số được chứa trong APS 716. Các kiểu này có thể bao gồm APS bộ lọc vòng lặp thích ứng (adaptive loop filter, ALF), APS ánh xạ độ chói với việc định tỷ lệ sắc độ (LMCS), và APS danh sách định tỷ lệ (Scaling). ALF là khối thích ứng dựa vào bộ lọc bao gồm chức năng truyền được điều khiển bởi các tham số biến đổi được và sử dụng phản hồi từ vòng lặp phản hồi để tinh lọc chức năng truyền. Ngoài ra, ALF được sử dụng để hiệu chỉnh các nhiễu lặp mã (ví dụ các lỗi chẳng hạn) xảy ra dưới dạng kết quả của việc lập mã dựa vào khối, như các nhiễu nhòe và vòng. Do đó, các tham số ALF có trong ALF APS có thể bao gồm các tham số được chọn bởi bộ mã hóa để khiến ALF loại bỏ nhiễu lặp mã dựa vào khối trong khi giải mã ở bộ giải mã. LMCS là quy trình được áp dụng dưới dạng một phần của quy trình giải mã ánh xạ các mẫu độ chói đến các giá trị cụ thể và trong một số trường hợp cũng áp dụng các hoạt động định tỷ lệ cho các giá trị của các mẫu sắc độ. Công cụ LMCS có thể tạo hình lại các thành phần độ chói dựa vào các ánh xạ đến các thành phần sắc độ tương ứng để giảm tỷ lệ biến dạng. Do đó, LMCS APS bao gồm các tham số được chọn bởi bộ mã hóa để khiến công cụ LMCS tạo hình lại các thành phần độ chói. Danh sách định tỷ lệ APS chứa các tham số công cụ lập mã được liên kết với các ma trận lượng tử hóa được sử dụng bởi các bộ lọc xác định. Do đó, APS 716 có thể chứa các tham số được sử dụng để áp dụng các bộ lọc khác nhau cho các lát được lập mã 727 trong khi kiểm tra sự tương thích ở HRD và/hoặc khi giải mã ở bộ giải mã.

Phần đầu hình ảnh 718 là cấu trúc cú pháp chứa các phần tử cú pháp áp dụng cho tất cả các lát 727 của hình ảnh được lập mã 725. Ví dụ, phần đầu hình ảnh 718 có thể chứa thông tin số đếm thứ tự hình ảnh, dữ liệu hình ảnh tham chiếu, dữ liệu liên quan trong các hình ảnh điểm truy cập ngẫu nhiên nội ảnh (intra-random access point, IRAP), dữ liệu có liên quan đến ứng dụng lọc cho hình ảnh 725, v.v.. PU có thể chứa chính xác một phần đầu hình ảnh 718 và chính xác một hình ảnh 725. Do đó, dòng bit 700 có thể bao gồm chính xác một phần đầu hình ảnh 718 trên mỗi hình ảnh 725. Phần đầu lát 717 chứa các tham số xác định cho mỗi lát 727 trong hình ảnh 725. Do đó, có thể có một phần đầu lát 717 trên lát 727 trong chuỗi video. Phần đầu lát 717 có thể chứa thông tin kiểu lát, thông tin bộ lọc, các trọng số của phép dự đoán, các điểm đầu vào ô, các tham số tách khối v.v.. Trong một số trường hợp, các phần tử cú pháp có thể giống nhau cho

tất cả các lát 727 trong hình ảnh 725. Để giảm sự dư thừa, phần đầu hình ảnh 718 và phần đầu lát 717 có thể chia sẻ các loại thông tin nhất định. Ví dụ, các tham số nhất định (ví dụ các tham số lọc) có thể có trong phần đầu hình ảnh 718 khi chúng áp dụng cho hình ảnh nguyên vẹn 725 hoặc có trong phần đầu lát 717 khi chúng áp dụng cho nhóm các lát 727 là tập hợp con của hình ảnh nguyên vẹn 725.

Dữ liệu hình ảnh 720 chứa dữ liệu video được mã hóa theo dự đoán liên ảnh và/hoặc dự đoán nội ảnh cũng như dữ liệu dư được biến đổi và được lượng tử hóa tương ứng. Ví dụ, dữ liệu hình ảnh 720 có thể bao gồm các lớp 723, các hình ảnh 725, và/hoặc các lát 727. Lớp 723 là tập hợp các đơn vị VCL NAL 745 chia sẻ một đặc trưng xác định (ví dụ độ phân giải chung, tốc độ khung hình, kích thước hình ảnh, v.v.) như được biểu thị bởi ID lớp, như `nuh_layer_id`, và các đơn vị non-VCL NAL có liên quan 741. Ví dụ, lớp 723 có thể bao gồm tập hợp các hình ảnh 725 chia sẻ cùng `nuh_layer_id`. Lớp 723 có thể gần như tương tự các lớp 631 và/hoặc 632. `nuh_layer_id` là phần tử cú pháp xác định mã định danh của lớp 723 bao gồm ít nhất một đơn vị NAL. Ví dụ, lớp chất lượng thấp nhất 723, được gọi là lớp cơ sở, có thể bao gồm giá trị thấp nhất của `nuh_layer_id` với các giá trị tăng của `nuh_layer_id` cho các lớp 723 có chất lượng cao hơn. Do đó, lớp dưới là lớp 723 có giá trị `nuh_layer_id` nhỏ hơn và lớp cao hơn là lớp 723 có giá trị `nuh_layer_id` lớn hơn.

Hình ảnh 725 là mảng gồm các mẫu độ chói và/hoặc mảng gồm các mẫu sắc độ tạo ra một khung hoặc trường của nó. Ví dụ, hình ảnh 725 là hình ảnh được lập mã có thể được xuất ra để hiển thị hoặc được sử dụng để hỗ trợ việc lập mã của (các) hình ảnh 725 khác để xuất ra. Hình ảnh 725 chứa một hoặc nhiều lát 727. Lát 727 có thể được xác định là số nguyên các ô hoàn chỉnh hoặc số nguyên các hàng đơn vị cây lập mã (CTU) hoàn chỉnh liên tiếp (ví dụ trong một ô) của hình ảnh 725 mà được chứa riêng trong một đơn vị NAL. Các lát 727 còn được chia thành các CTU và/hoặc các khối cây lập mã (các CTB). CTU là nhóm gồm các mẫu có kích thước định trước mà có thể được phân vùng bởi cây lập mã. CTB là tập hợp con của CTU và chứa các thành phần độ chói hoặc các thành phần sắc độ của CTU. Các CTU/ các CTB còn được chia thành các khối lập mã dựa vào các cây lập mã. Sau đó, các khối lập mã có thể được mã hóa/được giải mã theo cơ chế dự đoán.

Thông điệp SEI 719 là cấu trúc cú pháp có ngữ nghĩa xác định chuyển tải thông tin không cần thiết cho quy trình giải mã để xác định các giá trị của các mẫu trong các hình ảnh được giải mã. Ví dụ, các thông điệp SEI 719 có thể chứa dữ liệu hỗ trợ các quy trình HRD hoặc dữ liệu hỗ trợ khác không trực tiếp liên quan đến việc giải mã dòng bit 700 ở bộ giải mã. Tập hợp các thông điệp SEI 719 có thể được thực hiện dưới dạng thông điệp SEI lồng định tỷ lệ được. Thông điệp SEI lồng định tỷ lệ cung cấp cơ chế liên kết các thông điệp SEI 719 với các lớp 723 cụ thể. Thông điệp SEI lồng định tỷ lệ được là thông điệp chứa các thông điệp SEI được lồng-định tỷ lệ được. Thông điệp SEI được lồng-định tỷ lệ được là thông điệp SEI 719 tương ứng với một hoặc nhiều OLS hoặc một hoặc nhiều lớp 723. OLS là tập hợp các lớp 723, trong đó ít nhất một trong số các lớp 723 là lớp đầu ra. Do đó, thông điệp SEI lồng định tỷ lệ được có thể được xem là bao gồm tập hợp các thông điệp SEI được lồng-định tỷ lệ được hoặc được xem là bao gồm tập hợp các thông điệp SEI 719, tùy thuộc vào ngữ cảnh. Ngoài ra, thông điệp SEI lồng định tỷ lệ được chứa tập hợp các thông điệp SEI được lồng-định tỷ lệ được có cùng kiểu. Các thông điệp SEI 719 có thể bao gồm thông điệp BP SEI chứa các tham số HRD để khởi tạo HRD để quản lý CPB để kiểm tra các OLS tương ứng và/hoặc các lớp 723. Các thông điệp SEI 719 cũng có thể bao gồm thông điệp PT SEI chứa các tham số HRD để quản lý thông tin phân phối cho các AU ở CPB và/hoặc DPB để kiểm tra các OLS tương ứng và/hoặc các lớp 723. Các thông điệp SEI 719 cũng có thể bao gồm thông điệp DUI SEI chứa các tham số HRD để quản lý thông tin phân phối cho các DU ở CPB và/hoặc DPB để kiểm tra các OLS tương ứng và/hoặc các lớp 723.

Dòng bit 700 có thể được lập mã dưới dạng một trình tự của các đơn vị NAL. Đơn vị NAL là bộ phận chứa để chứa dữ liệu video và/hoặc cú pháp hỗ trợ. Đơn vị NAL có thể là đơn vị VCL NAL 745 hoặc đơn vị non-VCL NAL 741. Đơn vị VCL NAL 745 là đơn vị NAL được lập mã để chứa dữ liệu video. Cụ thể là, đơn vị VCL NAL 745 chứa lát 727 và phần đầu lát 717 có liên quan. Đơn vị non-VCL NAL 741 là đơn vị NAL chứa dữ liệu không phải video như cú pháp và/hoặc các tham số hỗ trợ việc giải mã dữ liệu video, hiệu suất của phép kiểm tra sự tương thích, hoặc các hoạt động khác. Các đơn vị non-VCL NAL 741 có thể bao gồm đơn vị NAL VPS, đơn vị NAL SPS, đơn vị NAL PPS, đơn vị NAL APS, đơn vị NAL phần đầu hình ảnh (picture header, PH), và đơn vị SEI NAL, lần lượt chứa VPS 711, SPS 713, PPS 715, APS 716, phần đầu hình ảnh 718,

và thông điệp SEI 719. Cần lưu ý rằng danh sách có trước của các đơn vị NAL chỉ là ví dụ và không phải là đầy đủ.

Mỗi đơn vị NAL được liên kết với mã định danh theo thời gian phần đầu đơn vị NAL cộng một ($\text{nuh_temporal_id_plus1}$) 731. $\text{nuh_temporal_id_plus1}$ 731 là mã định danh được báo hiệu biểu thị vị trí tương đối của đơn vị NAL tương ứng trong chuỗi video. Bộ giải mã và/hoặc HRD có thể xác định TemporalId cho đơn vị NAL tương ứng dựa vào giá trị $\text{nuh_temporal_id_plus1}$ 731. Cụ thể là, $\text{nuh_temporal_id_plus1}$ 731 được báo hiệu trong phần đầu của đơn vị NAL. TemporalId cho đơn vị NAL có thể được xác định bởi bộ giải mã/HRD bằng cách trừ một từ giá trị của $\text{nuh_temporal_id_plus1}$ 731. Do đó, giá trị của $\text{nuh_temporal_id_plus1}$ 731 không nên được thiết lập bằng zero do sẽ dẫn đến TemporalId có giá trị âm.

Ngoài ra, các thông điệp SEI 719 có thể được sử dụng làm các thông điệp SEI tiên tố và/hoặc các thông điệp SEI hậu tố. Thông điệp SEI tiên tố là thông điệp SEI 719 áp dụng cho một hoặc nhiều đơn vị NAL sao đó. Thông điệp SEI hậu tố là thông điệp SEI 719 áp dụng cho một hoặc nhiều đơn vị NAL đứng trước. Thông điệp SEI tiên tố có trong kiểu đơn vị SEI NAL tiên tố (PREFIX_SEI_NUT) 742 và thông điệp SEI hậu tố có trong kiểu đơn vị SEI NAL hậu tố (SUFFIX_SEI_NUT) 743. PREFIX_SEI_NUT 742 là đơn vị non-VCL NAL có tập hợp giá trị kiểu để biểu thị đơn vị non-VCL NAL chứa thông điệp SEI tiên tố. SUFFIX_SEI_NUT 743 là đơn vị non-VCL NAL có tập hợp giá trị kiểu để biểu thị đơn vị non-VCL NAL chứa thông điệp SEI hậu tố.

Như được lưu ý ở trên, các thông điệp SEI 719 có thể chứa các tham số được sử dụng bởi HRD hoạt động ở bộ mã hóa để kiểm tra dòng bit về sự tương thích với các chuẩn. Các thông điệp SEI 719 có thể có liên quan đến các hình ảnh khác nhau 725 và/hoặc các tổ hợp khác nhau của các lớp 723. Do đó, việc bảo đảm rằng thông điệp SEI thích hợp 719 được liên kết với các hình ảnh 725 và/hoặc các lớp 723 hợp quy trở thành thách thức trong các dòng bit đa lớp phức tạp. Ngoài ra, thông điệp SEI tiên tố có thể có trong dòng bit 700 trước khi đơn vị NAL thứ nhất được liên kết với thông điệp SEI tiên tố, trong khi thông điệp SEI hậu tố có thể có trong dòng bit 700 ngay sau khi đơn vị NAL thứ nhất được liên kết với thông điệp SEI hậu tố. Trong trường hợp thông điệp SEI 719 không được định vị đúng trong dòng bit 700 và/hoặc không được liên kết với lớp 723 và/hoặc hình ảnh 725 đúng, HRD có thể không thể kiểm tra lớp 723 và/hoặc hình

ảnh 725 về sự tương thích theo cách đúng đắn. Việc này có thể dẫn đến các sai số mã hóa gây ra bởi HRD và/hoặc các lỗi khi giải mã ở bộ giải mã. Ví dụ, HRD có thể lọc hình ảnh không đúng cách và/hoặc không phát hiện ra các xung đột về chuẩn. Ngoài ra, bộ giải mã có thể không phát hiện ra truyền dẫn liên quan đến các lỗi lập mã và/hoặc trả không đúng cách chỉ báo về lỗi lập mã liên quan đến khi không tồn tại các lỗi như vậy.

Dòng bit 700 được cải biến để bảo đảm các thông điệp SEI 719 được liên kết đúng với các lớp tương ứng 723, các hình ảnh 725, các lát 727, và/hoặc các đơn vị NAL. Như được mô tả đối với FIG.6, các dòng bit đa lớp có thể tổ chức các hình ảnh và các tham số có liên quan thành các AU. AU là tập hợp các hình ảnh được lập mã có trong các lớp khác nhau và được liên kết với cùng thời gian đầu ra. Trong dòng bit 700, mỗi thông điệp SEI 719 được định vị trong cùng AU dưới dạng hình ảnh thứ nhất 725 được liên kết với thông điệp SEI 719. Ngoài ra, thông điệp SEI 719 được gán TemporalId. TemporalId của thông điệp SEI 719 được ràng buộc bằng TemporalId của AU chứa thông điệp SEI 719. Trong một ví dụ cụ thể, mỗi hình ảnh 725 trong AU chia sẻ cùng giá trị của TemporalId và do đó chia sẻ cùng giá trị của `nuh_temporal_id_plus1` 731. Do đó, `nuh_temporal_id_plus1` 731 của thông điệp SEI 719 là giống như `nuh_temporal_id_plus1` 731 của mỗi hình ảnh 725 tương ứng với hình ảnh mà thông điệp SEI 719 áp dụng.

Nói cách khác, các hình ảnh 725 có trong các đơn vị VCL NAL 745 và các tham số có trong các đơn vị non-VCL NAL 741. Khi đơn vị non-VCL NAL 741 là đơn vị SEI NAL thuộc kiểu `PREFIX_SEI_NUT` 742 hoặc `SUFFIX_SEI_NUT` 743 chứa thông điệp SEI 719, TemporalId/ `nuh_temporal_id_plus1` 731 của đơn vị non-VCL NAL 741 được ràng buộc bằng TemporalId/ `nuh_temporal_id_plus1` 731 của AU chứa đơn vị non-VCL NAL 741. Phương pháp này bảo đảm rằng các thông điệp SEI 719 được liên kết đúng với các hình ảnh tương ứng 725 trong các AU. Do đó, có thể tránh được các lỗi khác nhau. Kết quả là, chức năng của bộ mã hóa và bộ giải mã tăng. Ngoài ra, hiệu quả lập mã có thể tăng lên, giảm việc sử dụng tài nguyên báo hiệu bộ xử lý, bộ nhớ và/hoặc mạng ở cả bộ mã hóa lẫn bộ giải mã.

Tiếp theo, thông tin có trước được mô tả chi tiết hơn dưới đây. Việc lập mã video phân lớp còn được gọi là lập mã video phân cấp được hoặc lập mã video có khả năng phân cấp. Khả năng phân cấp trong lập mã video có thể được hỗ trợ bằng cách sử dụng

kỹ thuật lập mã đa lớp. Dòng bit đa lớp bao gồm lớp cơ sở (BL) và một hoặc nhiều lớp nâng cao (enhancement layer, EL). Ví dụ về khả năng phân cấp bao gồm khả năng phân cấp theo không gian, khả năng phân cấp chất lượng / tỷ lệ tín hiệu trên nhiễu (signal to noise ratio, SNR), khả năng phân cấp đa góc nhìn, khả năng phân cấp tốc độ khung hình, v.v.. Khi kỹ thuật lập mã đa lớp được sử dụng, hình ảnh hoặc một phần của nó có thể được lập mã mà không sử dụng hình ảnh tham chiếu (dự đoán nội ảnh), có thể được lập mã bằng cách tham chiếu các hình ảnh tham chiếu trong cùng lớp (dự đoán liên ảnh), và/hoặc có thể được lập mã bằng cách tham chiếu các hình ảnh tham chiếu nằm trong (các) lớp khác (phép dự đoán liên-lớp). Hình ảnh tham chiếu được sử dụng cho phép dự đoán liên-lớp của hình ảnh hiện thời được gọi là hình ảnh tham chiếu liên-lớp (ILRP). FIG.6 minh họa ví dụ về lập mã đa lớp cho khả năng phân cấp theo không gian trong đó các hình ảnh trong các lớp khác nhau có các độ phân giải khác nhau.

Một số dòng lập mã video cung cấp sự hỗ trợ cho khả năng phân cấp trong (các) hồ sơ riêng biệt từ (các) hồ sơ dùng cho việc lập mã lớp đơn. Lập mã video phân cấp được (Scalable video coding, SVC) là phần mở rộng phân cấp được của lập mã video tiên tiến (advanced video coding, AVC) cung cấp sự hỗ trợ cho khả năng phân cấp theo không gian, khả năng phân cấp theo thời gian, và khả năng phân cấp chất lượng. Đối với SVC, một cờ được báo hiệu trong mỗi khối macro (MB) trong các hình ảnh EL để biểu thị việc EL MB có được dự đoán hay không bằng cách sử dụng khối sắp xếp từ lớp dưới. Phép dự đoán từ khối sắp xếp có thể bao gồm mẫu hình bề mặt, các vectơ chuyển động, và/hoặc các chế độ lập mã. Các phương án thực hiện của SVC có thể không trực tiếp sử dụng lại các phương án thực hiện AVC không được cải biến trong thiết kế của chúng. Cú pháp khối macro SVC EL và quy trình giải mã khác với cú pháp AVC và quy trình giải mã.

HEVC phân cấp được (SHVC) là phần mở rộng của HEVC cung cấp sự hỗ trợ cho các khả năng phân cấp theo không gian và chất lượng. Đa điểm HEVC (MV-HEVC) là phần mở rộng của HEVC cung cấp sự hỗ trợ cho khả năng phân cấp đa góc nhìn. 3D HEVC (3D-HEVC) là phần mở rộng của HEVC cung cấp sự hỗ trợ cho lập mã video 3D tiên tiến hơn và có hiệu quả hơn so với MV-HEVC. Khả năng phân cấp theo thời gian có thể bao gồm dưới dạng một phần tích hợp của bộ lập-giải mã HEVC một lớp. Trong phần mở rộng đa lớp của HEVC, các hình ảnh được giải mã được sử dụng cho

phép dự đoán liên-lớp chỉ đến từ cùng AU và được xử lý dưới dạng các hình ảnh tham chiếu lâu dài (long-term reference picture, LTRP). Các hình ảnh này được gán các chỉ số tham chiếu trong (các) danh sách hình ảnh tham chiếu cùng với các hình ảnh tham chiếu theo thời gian khác trong lớp hiện thời. Phép dự đoán liên-lớp (ILP) đạt được ở cấp độ đơn vị dự đoán nhờ đặt giá trị của chỉ số tham chiếu liên quan đến (các) hình ảnh tham chiếu liên-lớp trong (các) danh sách hình ảnh tham chiếu. Khả năng phân cấp theo không gian lấy mẫu lại hình ảnh tham chiếu hoặc một phần của nó khi ILRP có độ phân giải theo không gian khác với hình ảnh hiện thời được mã hóa hoặc được giải mã. Việc lấy mẫu lại hình ảnh tham chiếu có thể được thực hiện ở cấp độ hình ảnh hoặc cấp độ khối lập mã.

Chuẩn VVC cũng có thể hỗ trợ việc lập mã video phân lớp. Dòng bit VVC có thể bao gồm nhiều lớp. Các lớp này có thể đều độc lập với nhau với nhau. Ví dụ, mỗi lớp có thể được lập mã mà không sử dụng phép dự đoán liên-lớp. Trong trường hợp này, các lớp còn được gọi là các lớp phát đồng thời. Trong một số trường hợp, một số lớp được lập mã bằng cách sử dụng ILP. Còn trong VPS có thể biểu thị việc các lớp này có phải là các lớp phát đồng thời hay không hoặc việc một số lớp có sử dụng ILP hay không. Khi một số lớp sử dụng ILP, mối quan hệ phụ thuộc lớp trong số các lớp cũng được báo hiệu trong VPS. Không giống SHVC và MV-HEVC, VVC có thể không xác định các OLS. OLS bao gồm tập hợp xác định gồm các lớp, trong đó một hoặc nhiều lớp trong tập hợp các lớp được xác định là các lớp đầu ra. Lớp đầu ra là một lớp của OLS được xuất ra. Trong VVC theo một số phương án thực hiện, chỉ một lớp có thể được chọn để giải mã và xuất ra khi các lớp các lớp phát đồng thời. Trong VVC theo một số phương án thực hiện, toàn bộ dòng bit bao gồm tất cả các lớp được xác định là được giải mã khi lớp bất kỳ sử dụng ILP. Ngoài ra, các lớp nhất định trong số các lớp này được xác định là các lớp đầu ra. Các lớp đầu ra có thể được biểu thị là chỉ một lớp cao nhất, tất cả các lớp, hoặc lớp cao nhất cộng tập hợp các lớp dưới được biểu thị.

Các khía cạnh nêu trên chứa các vấn đề nhất định. Ví dụ, các giá trị `nuh_layer_id` cho SPS, PPS, và các đơn vị NAL APS có thể không được ràng buộc đúng. Ngoài ra, giá trị `TemporalId` cho các đơn vị SEI NAL cũng có thể không được ràng buộc đúng. Ngoài ra, việc thiết lập `NoOutputOfPriorPicsFlag` có thể không được xác định đúng khi việc lấy mẫu lại hình ảnh tham chiếu được cho phép và các hình ảnh trong CLVS có các

độ phân giải theo không gian khác nhau. Ngoài ra, trong một số hệ thống lập mã video, các thông điệp SEI hậu tố không thể được chứa trong thông điệp SEI lồng định tỷ lệ được. Trong một ví dụ khác, chu kỳ đệm, việc định thời hình ảnh, và các thông điệp SEI thông tin đơn vị giải mã có thể bao gồm sự phụ thuộc phân tích vào VPS và/hoặc SPS.

Nói chung, sáng chế mô tả phương pháp cải tiến việc lập mã video. Các phần mô tả kỹ thuật này dựa vào chuẩn VVC. Tuy nhiên, các kỹ thuật này cũng áp dụng được cho việc lập mã video theo lớp dựa vào các thông số lập-giải mã video khác.

Một hoặc nhiều vấn đề nêu trên có thể được giải quyết như sau. Các giá trị `nuh_layer_id` cho SPS, PPS, và các đơn vị NAL APS được ràng buộc đúng ở đây. Giá trị `TemporalId` cho các đơn vị SEI NAL được ràng buộc đúng ở đây. Việc thiết lập `NoOutputOfPriorPicsFlag` được xác định đúng khi việc lấy mẫu lại hình ảnh tham chiếu được cho phép và các hình ảnh bên trong CLVS có các độ phân giải theo không gian khác nhau. Các thông điệp SEI hậu tố được cho phép chứa trong thông điệp SEI lồng định tỷ lệ được. Sự phụ thuộc phân tích của các thông điệp BP, PT, và DUI SEI trên VPS hoặc SPS có thể được loại bỏ bằng cách lặp lại phần tử cú pháp `decoding_unit_hrd_params_present_flag` trong cú pháp thông điệp BP SEI, các phần tử cú pháp `decoding_unit_hrd_params_present_flag` và `decoding_unit_cpb_params_in_pic_timing_sei_flag` trong cú pháp thông điệp PT SEI, và phần tử cú pháp `decoding_unit_cpb_params_in_pic_timing_sei_flag` trong thông điệp DUI SEI.

Một phương án thực hiện làm ví dụ của cơ chế nêu trên như sau. Ví dụ, ngữ nghĩa của đơn vị NAL chung như sau.

`nuh_temporal_id_plus1` trừ 1 xác định mã định danh theo thời gian cho đơn vị NAL. giá trị của `nuh_temporal_id_plus1` sẽ không bằng zero. Biến `TemporalId` có thể được suy ra như sau:

$$\text{TemporalId} = \text{nuh_temporal_id_plus1} - 1$$

Khi `nal_unit_type` nằm trong khoảng từ `IDR_W_RADL` đến `RSV_IRAP_13`, bao gồm cả hai giá trị này, `TemporalId` sẽ bằng zero. Khi `nal_unit_type` bằng `STSA_NUT`, `TemporalId` sẽ không bằng zero.

Giá trị của TemporalId nên giống nhau đối với tất cả các đơn vị VCL NAL của đơn vị truy cập. Giá trị của TemporalId của hình ảnh được lập mã, đơn vị truy cập lớp, hoặc đơn vị truy cập có thể là giá trị của TemporalId của các đơn vị VCL NAL của hình ảnh được lập mã, đơn vị truy cập lớp, hoặc đơn vị truy cập. Giá trị của TemporalId của biểu diễn lớp con có thể là giá trị lớn nhất của TemporalId của tất cả các đơn vị VCL NAL trong biểu diễn lớp con này.

Giá trị của TemporalId cho các đơn vị non-VCL NAL được ràng buộc như sau. nếu nal_unit_type bằng DPS_NUT, VPS_NUT, hoặc SPS_NUT, TemporalId bằng zero và TemporalId của đơn vị truy cập chứa đơn vị NAL sẽ bằng zero. Theo cách khác, nếu nal_unit_type bằng EOS_NUT hoặc EOB_NUT, TemporalId sẽ bằng zero. Theo cách khác, nếu nal_unit_type bằng AUD_NUT, FD_NUT, PREFIX_SEI_NUT, hoặc SUFFIX_SEI_NUT, TemporalId sẽ bằng TemporalId của đơn vị truy cập chứa đơn vị NAL. Theo cách khác, khi nal_unit_type bằng PPS_NUT hoặc APS_NUT, TemporalId sẽ lớn hơn hoặc bằng TemporalId của đơn vị truy cập chứa đơn vị NAL. Khi đơn vị NAL là đơn vị non-VCL NAL, giá trị của TemporalId sẽ bằng giá trị nhỏ nhất trong số các giá trị TemporalId của tất cả các đơn vị truy cập mà đơn vị non-VCL NAL áp dụng. Khi nal_unit_type bằng PPS_NUT hoặc APS_NUT, TemporalId có thể lớn hơn hoặc bằng TemporalId của đơn vị truy cập đang bao gồm. Điều này là vì tất cả các PPS và các APS có thể đều có trong phần bắt đầu của dòng bit. Ngoài ra, hình ảnh thứ nhất được lập mã có TemporalId bằng zero.

Một ví dụ về ngữ nghĩa tập hợp tham số trình tự RBSP như sau. RBSP SPS sẽ khả dụng cho quy trình giải mã trước khi được tham chiếu. SPS có thể có trong ít nhất một đơn vị truy cập với TemporalId bằng zero hoặc được cung cấp bởi cơ chế bên ngoài. Đơn vị NAL SPS chứa SPS có thể được ràng buộc để có nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các đơn vị NAL PPS liên quan đến SPS.

Một ví dụ về ngữ nghĩa tập hợp tham số hình ảnh RBSP như sau. RBSP PPS sẽ khả dụng cho quy trình giải mã trước khi được tham chiếu. PPS sẽ có trong ít nhất một đơn vị truy cập với TemporalId nhỏ hơn hoặc bằng TemporalId của đơn vị NAL PPS hoặc được cung cấp qua cơ chế bên ngoài. Đơn vị NAL PPS chứa RBSP PPS sẽ có nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các đơn vị NAL lát được lập mã liên quan đến PPS.

Một ví dụ về ngữ nghĩa tập hợp tham số thích ứng như sau. Mỗi RBSP APS sẽ khả dụng cho quy trình giải mã trước khi được tham chiếu. APS cũng sẽ có trong ít nhất một đơn vị truy cập với TemporalId nhỏ hơn hoặc bằng TemporalId của lát được lập mã đơn vị NAL liên quan đến APS hoặc được cung cấp nhờ cơ chế bên ngoài. Đơn vị NAL APS được cho phép chia sẻ với các hình ảnh/các lát của nhiều lớp. `nuh_layer_id` của đơn vị NAL APS sẽ bằng giá trị `nuh_layer_id` thấp nhất của lát được lập mã các đơn vị NAL liên quan đến APS đơn vị NAL. Theo cách khác, đơn vị NAL APS có thể không được chia sẻ bởi các hình ảnh/các lát của nhiều lớp. `nuh_layer_id` của đơn vị NAL APS sẽ bằng `nuh_layer_id` của các lát liên quan đến APS.

Trong một ví dụ, việc loại bỏ các hình ảnh ra khỏi DPB trước khi giải mã hình ảnh hiện thời được trình bày như sau. Việc loại bỏ các hình ảnh ra khỏi DPB trước khi giải mã hình ảnh hiện thời (nhưng sau khi phân tích phần đầu lát của lát thứ nhất của hình ảnh hiện thời) có thể diễn ra ở thời gian loại bỏ CPB của đơn vị giải mã thứ nhất của đơn vị truy cập *n* (chứa hình ảnh hiện thời). Việc này diễn ra như sau. Quy trình giải mã cho việc xây dựng danh sách hình ảnh tham chiếu được gọi và quy trình giải mã cho việc đánh dấu hình ảnh tham chiếu được gọi.

Khi hình ảnh hiện thời là hình ảnh phần bắt đầu chuỗi video lớp được lập mã (CLVSS) không phải là zero hình ảnh, các bước theo thứ tự sau được áp dụng. Biến `NoOutputOfPriorPicsFlag` được suy ra cho bộ giải mã dưới kiểm tra như sau. Nếu giá trị của `pic_width_max_in_luma_samples`, `pic_height_max_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_flag`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ Htid ]` được suy ra từ SPS lần lượt khác với giá trị của `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `chroma_format_idc`, `separate_colour_plane_flag`, `bit_depth_luma_minus8`, `bit_depth_chroma_minus8` hoặc `sps_max_dec_pic_buffering_minus1[ Htid ]`, được suy ra từ SPS được tham chiếu bởi hình ảnh trước đó, `NoOutputOfPriorPicsFlag` có thể được thiết lập bằng một bởi bộ giải mã dưới kiểm tra, bất kể giá trị của `no_output_of_prior_pics_flag`. Cần lưu ý rằng, mặc dù việc thiết lập `NoOutputOfPriorPicsFlag` bằng `no_output_of_prior_pics_flag` có thể được ưu tiên dưới các điều kiện này, nhưng bộ giải mã dưới kiểm tra được cho phép để

thiết lập NoOutputOfPriorPicsFlag bằng một trong trường hợp này. Theo cách khác, NoOutputOfPriorPicsFlag có thể thiết lập bằng no_output_of_prior_pics_flag.

Giá trị của NoOutputOfPriorPicsFlag được suy ra cho bộ giải mã dưới kiểm tra được áp dụng cho HRD, sao cho khi giá trị của NoOutputOfPriorPicsFlag bằng 1, tất cả các bộ đệm lưu trữ hình ảnh trong DPB được làm rỗng mà không xuất ra các hình ảnh mà chúng chứa, và lượng lấp đầy DPB được thiết lập bằng zero. Khi cả hai điều kiện nêu trên đều đúng đối với các hình ảnh k bất kỳ trong DPB, tất cả các hình ảnh k trong DPB được loại ra khỏi DPB. Hình ảnh k được đánh dấu là không được sử dụng để tham chiếu, và hình ảnh k có PictureOutputFlag bằng zero hoặc thời gian đầu ra DPB tương ứng nhỏ hơn hoặc bằng thời gian loại bỏ CPB của đơn vị giải mã thứ nhất (được thể hiện là đơn vị giải mã m) của hình ảnh hiện thời n. Việc này có thể xảy ra khi DpbOutputTime [k] nhỏ hơn hoặc bằng DuCpbRemovalTime[m]. Đối với mỗi hình ảnh được loại ra khỏi DPB, lượng lấp đầy DPB giảm một.

Trong một ví dụ, việc xuất ra và loại bỏ các hình ảnh ra khỏi DPB được trình bày như sau. Việc xuất ra và loại bỏ các hình ảnh ra khỏi DPB trước khi giải mã hình ảnh hiện thời (nhưng sau khi phân tích phần đầu lát của lát thứ nhất của hình ảnh hiện thời) có thể diễn ra khi đơn vị giải mã thứ nhất của đơn vị truy cập chứa hình ảnh hiện thời được loại ra khỏi CPB và tiến hành như sau. Quy trình giải mã cho việc xây dựng danh sách hình ảnh tham chiếu và quy trình giải mã cho việc đánh dấu hình ảnh tham chiếu được gọi.

Nếu hình ảnh hiện thời là hình ảnh CLVSS không phải là hình ảnh zero, các bước theo thứ tự sau được áp dụng. biến NoOutputOfPriorPicsFlag có thể được suy ra cho bộ giải mã dưới kiểm tra như sau. Nếu giá trị của pic_width_max_in_luma_samples, pic_height_max_in_luma_samples, chroma_format_idc, separate_colour_plane_flag, bit_depth_luma_minus8, bit_depth_chroma_minus8 hoặc sps_max_dec_pic_buffering_minus1[Htid] được suy ra từ SPS lần lượt khác giá trị của pic_width_in_luma_samples, pic_height_in_luma_samples, chroma_format_idc, separate_colour_plane_flag, bit_depth_luma_minus8, bit_depth_chroma_minus8 hoặc sps_max_dec_pic_buffering_minus1[Htid], được suy ra từ SPS được tham chiếu bởi hình ảnh trước đó, NoOutputOfPriorPicsFlag có thể được thiết lập bằng một bởi bộ giải mã dưới kiểm tra, bất kể giá trị của no_output_of_prior_pics_flag. Cần lưu ý rằng mặc

dù việc thiết lập NoOutputOfPriorPicsFlag bằng no_output_of_prior_pics_flag được ưu tiên dưới các điều kiện này, nhưng bộ giải mã dưới kiểm tra có thể thiết lập NoOutputOfPriorPicsFlag bằng một trong trường hợp này. Theo cách khác, NoOutputOfPriorPicsFlag có thể được thiết lập bằng no_output_of_prior_pics_flag.

Giá trị của NoOutputOfPriorPicsFlag được suy ra cho bộ giải mã dưới kiểm tra có thể được áp dụng cho HRD như sau. Nếu NoOutputOfPriorPicsFlag bằng một, tất cả các bộ đệm lưu trữ hình ảnh trong DPB được làm rỗng mà không xuất ra các hình ảnh mà chúng chứa và lượng lấp đầy DPB được thiết lập bằng zero. Theo cách khác (NoOutputOfPriorPicsFlag bằng zero), tất cả các bộ đệm lưu trữ hình ảnh chứa hình ảnh được đánh dấu là không cần thiết cho đầu ra và không được sử dụng để tham chiếu được làm rỗng (mà không xuất ra) và tất cả các bộ đệm lưu trữ hình ảnh không rỗng trong DPB được làm rỗng bằng cách gọi lặp đi lặp lại quy trình bumping và lượng lấp đầy DPB được thiết lập bằng zero.

Theo cách khác (hình ảnh hiện thời không phải là hình ảnh CLVSS), tất cả các bộ đệm lưu trữ hình ảnh chứa hình ảnh được đánh dấu là không cần thiết cho đầu ra và không được sử dụng để tham chiếu được làm rỗng (mà không xuất ra). Đối với mỗi bộ đệm lưu trữ hình ảnh được làm rỗng, lượng lấp đầy DPB được giảm thêm một. Khi một hoặc nhiều điều kiện nêu trên đúng, quy trình bumping được gọi lặp đi lặp lại trong khi giảm thêm lượng lấp đầy DPB bằng một đối với mỗi bộ đệm lưu trữ hình ảnh bổ sung được làm rỗng cho đến khi không còn điều kiện nào trong số các điều kiện nêu trên đúng. Điều kiện là số lượng hình ảnh trong DPB được đánh dấu khi cần cho đầu ra lớn hơn sps_max_num_reorder_pics[Htid]. Một điều kiện khác là sps_max_latency_increase_plus1[Htid] không bằng zero và có ít nhất một hình ảnh trong DPB được đánh dấu khi cần cho đầu ra cho biến PicLatencyCount được liên kết lớn hơn hoặc bằng SpsMaxLatencyPictures[Htid]. Một điều kiện khác là số lượng hình ảnh trong DPB lớn hơn hoặc bằng SubDpbSize[Htid].

Ví dụ về cú pháp thông điệp SEI chung như sau.

sei_payload(payloadType, payloadSize) {	Khóa mô tả
---	------------

if(nal_unit_type == PREFIX_SEI_NUT)	
if(payloadType == 0)	
buffering_period(payloadSize)	
else if(payloadType == 1)	
pic_timing(payloadSize)	
else if(payloadType == 3)	
filler_payload(payloadSize)	
else if(payloadType == 130)	
decoding_unit_info(payloadSize)	
else if(payloadType == 133)	
scalable_nesting(payloadSize)	
else if(payloadType == 145)	
dependent_rap_indication(payloadSize) // được xác định trong ITU-T H.SEI ISO/IEC 23002-7.	
else if(payloadType == 168)	
frame_field_info(payloadSize)	
else	
reserved_sei_message(payloadSize)	
else /* nal_unit_type == SUFFIX_SEI_NUT */	
if(payloadType == 3)	

filler_payload(payloadSize)	
if(payloadType == 132)	
decoded_picture_hash(payloadSize) // được xác định trong ITU-T H.SEI ISO/IEC 23002-7.	
else if(payloadType == 133)	
scalable_nesting(payloadSize)	
else	
reserved_sei_message(payloadSize)	
if(more_data_in_payload()) {	
if(payload_extension_present())	
reserved_payload_extension_data	u(v)
payload_bit_equal_to_one /* equal to 1 */	f(1)
while(!byte_aligned())	
payload_bit_equal_to_zero /* equal to 0 */	f(1)
}	
}	

Ví dụ về cú pháp thông điệp SEI lồng định tỷ lệ được như sau.

scalable_nesting(payloadSize) {	Khóa mô tả
nesting_ols_flag	u(1)

if(nesting_ols_flag) {	
nesting_num_olss_minus1	ue(v)
for(i = 0; i <= nesting_num_olss_minus1; i++) {	
nesting_ols_idx_delta_minus1[i]	ue(v)
if(NumLayersInOls[NestingOlsIdx[i]] > 1) {	
nesting_num_ols_layers_minus1[i]	ue(v)
for(j = 0; j <= nesting_num_ols_layers_minus1[i]; j++)	
nesting_ols_layer_idx_delta_minus1[i][j]	ue(v)
}	
}	
} else {	
nesting_all_layers_flag	u(1)
if(!nesting_all_layers_flag) {	
nesting_num_layers_minus1	ue(v)
for(i = 1; i <= nesting_num_layers_minus1; i++)	
nesting_layer_id[i]	u(6)
}	
}	
nesting_num_seis_minus1	ue(v)
while(!byte_aligned())	

nesting_zero_bit /* equal to 0 */	u(1)
for(i = 0; i <= nesting_num_seis_minus1; i++)	
sei_message()	
}	

Ví dụ về ngữ nghĩa thông điệp SEI lồng định tỷ lệ được như sau. Thông điệp SEI lồng định tỷ lệ được cung cấp cơ chế để liên kết các thông điệp SEI với các lớp cụ thể trong ngữ cảnh của các OLS cụ thể hoặc với các lớp cụ thể không trong ngữ cảnh của OLS. Thông điệp SEI lồng định tỷ lệ được chứa một hoặc nhiều thông điệp SEI. Các thông điệp SEI chứa trong thông điệp SEI lồng định tỷ lệ được còn được gọi là các thông điệp SEI được lồng-định tỷ lệ được. Sự tương thích dòng bit có thể yêu cầu rằng các giới hạn sau đây áp dụng khi các thông điệp SEI được chứa trong thông điệp SEI lồng định tỷ lệ được.

Một thông điệp SEI có payloadType bằng một trăm ba mươi hai (bấm hình ảnh được giải mã) hoặc một trăm ba mươi ba (lồng định tỷ lệ được) không được chứa trong thông điệp SEI lồng định tỷ lệ được. Khi thông điệp SEI lồng định tỷ lệ được chứa thông điệp SEI chu kỳ đệm, định thời hình ảnh, hoặc thông tin đơn vị giải mã, thông điệp SEI lồng định tỷ lệ được sẽ không chứa bất kỳ thông điệp SEI nào khác có payloadType không bằng zero (chu kỳ đệm), một (định thời hình ảnh), hoặc một trăm ba mươi (thông tin đơn vị giải mã).

Sự tương thích dòng bit cũng có thể yêu cầu rằng các giới hạn sau đây áp dụng trên giá trị của nal_unit_type của đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được. Khi thông điệp SEI lồng định tỷ lệ được chứa thông điệp SEI có payloadType bằng zero (chu kỳ đệm), một (định thời hình ảnh), một trăm ba mươi (thông tin đơn vị giải mã), một trăm bốn mươi lăm (chỉ báo RAP phụ thuộc), hoặc một trăm sáu mươi tám (thông tin trường khung), đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được cần có nal_unit_type được thiết lập bằng PREFIX_SEI_NUT. Khi thông điệp SEI lồng định tỷ lệ được chứa thông điệp SEI có payloadType bằng một trăm ba mươi hai (bấm

hình ảnh được giải mã), đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được cần có nal_unit_type được thiết lập bằng SUFFIX_SEI_NUT.

esting_ols_flag có thể được thiết lập bằng một để xác định rằng các thông điệp SEI được lồng-định tỷ lệ được áp dụng cho các lớp cụ thể trong ngữ cảnh của các OLS cụ thể. nesting_ols_flag có thể được thiết lập bằng zero để xác định rằng các thông điệp SEI được lồng-định tỷ lệ được thường áp dụng (ví dụ không có trong ngữ cảnh của OLS) cho các lớp cụ thể.

Sự tương thích dòng bit có thể yêu cầu rằng các giới hạn sau đây được áp dụng cho giá trị của nesting_ols_flag. Khi thông điệp SEI lồng định tỷ lệ được chứa thông điệp SEI có payloadType bằng zero (chu kỳ đệm), một (định thời hình ảnh), hoặc một trăm ba mươi (thông tin đơn vị giải mã), giá trị của nesting_ols_flag sẽ bằng một. Khi thông điệp SEI lồng định tỷ lệ được chứa thông điệp SEI có payloadType bằng một giá trị trong VclAssociatedSeiList, giá trị của nesting_ols_flag sẽ bằng zero.

nesting_num_olss_minus1 cộng một xác định số lượng OLS mà các thông điệp SEI được lồng-định tỷ lệ được áp dụng. Giá trị của nesting_num_olss_minus1 sẽ nằm trong khoảng từ zero đến TotalNumOlss - 1, bao gồm cả hai giá trị này. nesting_ols_idx_delta_minus1[i] được sử dụng để suy ra biến NestingOlsIdx[i] xác định chỉ số OLS của OLS thứ i cho các thông điệp SEI được lồng-định tỷ lệ được áp dụng khi nesting_ols_flag bằng một. Giá trị của nesting_ols_idx_delta_minus1[i] sẽ nằm trong khoảng từ zero đến TotalNumOlss - 2, bao gồm cả hai giá trị này. Biến NestingOlsIdx[i] có thể được suy ra như sau:

```
if( i == 0 )
```

```
    NestingOlsIdx[ i ] = nesting_ols_idx_delta_minus1[ i ]
```

```
else
```

```
    NestingOlsIdx[ i ] = NestingOlsIdx[ i - 1 ] +
```

```
    nesting_ols_idx_delta_minus1[ i ] + 1
```

nesting_num_ols_layers_minus1[i] cộng một xác định số lượng lớp mà các thông điệp SEI được lồng-định tỷ lệ được áp dụng trong ngữ cảnh của OLS thứ NestingOlsIdx[i]. Giá trị của nesting_num_ols_layers_minus1[i] sẽ nằm trong khoảng từ zero đến NumLayersInOls[NestingOlsIdx[i]] - 1, bao gồm cả hai giá trị này.

`nesting_ols_layer_idx_delta_minus1[ i ][ j ]` được sử dụng để suy ra biến `NestingOlsLayerIdx[ i ][ j ]` xác định chỉ số lớp OLS của lớp thứ `j` mà các thông điệp SEI được lồng-định tỷ lệ được áp dụng trong ngữ cảnh của OLS thứ `NestingOlsIdx[ i ]` khi `nesting_ols_flag` bằng một. Giá trị của `nesting_ols_layer_idx_delta_minus1[ i ]` sẽ nằm trong khoảng từ zero đến `NumLayersInOls[ nestingOlsIdx[ i ] ] - hai`, bao gồm cả hai giá trị này.

Biến `NestingOlsLayerIdx[ i ][ j ]` có thể được suy ra như sau:

if(`j == 0`)

`NestingOlsLayerIdx[ i ][ j ] = nesting_ols_layer_idx_delta_minus1[ i ][ j ]`

else

`NestingOlsLayerIdx[ i ][ j ] = NestingOlsLayerIdx[ i ][ j - 1 ] +`

`nesting_ols_layer_idx_delta_minus1[ i ][ j ] + 1`

Giá trị thấp nhất trong số tất cả các giá trị của `LayerIdInOls[ NestingOlsIdx[ i ] ][ NestingOlsLayerIdx[ i ][ 0 ] ]` đối với `i` nằm trong khoảng từ zero đến `nesting_num_olss_minus1`, bao gồm cả hai giá trị này, nên bằng `nuh_layer_id` của đơn vị SEI NAL hiện thời (ví dụ đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được). `nesting_all_layers_flag` có thể được thiết lập bằng một để xác định rằng các thông điệp SEI được lồng-định tỷ lệ được thường áp dụng cho tất cả các lớp có `nuh_layer_id` lớn hơn hoặc bằng `nuh_layer_id` của đơn vị SEI NAL hiện thời. `nesting_all_layers_flag` có thể được thiết lập bằng zero để xác định rằng các thông điệp SEI được lồng-định tỷ lệ được có thể hoặc có thể thường không áp dụng cho tất cả các lớp mà có `nuh_layer_id` lớn hơn hoặc bằng `nuh_layer_id` của đơn vị SEI NAL hiện thời.

`nesting_num_layers_minus1` cộng một xác định số lượng lớp mà các thông điệp SEI được lồng-định tỷ lệ được thường áp dụng. Giá trị của `nesting_num_layers_minus1` sẽ nằm trong khoảng từ zero đến `vps_max_layers_minus1 - GeneralLayerIdx[ nuh_layer_id ]`, bao gồm cả hai giá trị này, trong đó `nuh_layer_id` là `nuh_layer_id` của đơn vị SEI NAL hiện thời. `nesting_layer_id[ i ]` xác định giá trị `nuh_layer_id` của lớp thứ `i` mà các thông điệp SEI được lồng-định tỷ lệ được thường áp dụng khi `nesting_all_layers_flag` bằng zero. Giá trị của `nesting_layer_id[ i ]` nên lớn hơn `nuh_layer_id`, trong đó `nuh_layer_id` là `nuh_layer_id` của đơn vị SEI NAL hiện thời.

Khi `nesting_ols_flag` bằng một, biến `NestingNumLayers`, xác định số lượng lớp mà các thông điệp SEI được lồng-định tỷ lệ được thường áp dụng, và danh sách `NestingLayerId[ i ]` đối với `i` nằm trong khoảng từ zero đến `NestingNumLayers - 1`, bao gồm cả hai giá trị này, xác định danh sách của giá trị `nuh_layer_id` của các lớp mà các thông điệp SEI được lồng-định tỷ lệ được thường áp dụng, được suy ra như sau, trong đó `nuh_layer_id` là `nuh_layer_id` của đơn vị SEI NAL hiện thời:

```

if( nesting_all_layers_flag ) {
    NestingNumLayers =
vps_max_layers_minus1 + 1 - GeneralLayerIdx[ nuh_layer_id ]
    for( i = 0; i < NestingNumLayers; i ++ )
        NestingLayerId[ i ] = vps_layer_id[ GeneralLayerIdx[ nuh_layer_id ] + i ]
(D-2)
} else {
    NestingNumLayers = nesting_num_layers_minus1 + 1
    for( i = 0; i < NestingNumLayers; i ++ )
        NestingLayerId[ i ] = ( i == 0 ) ? nuh_layer_id : nesting_layer_id[ i ]
}

```

`nesting_num_seis_minus1` cộng một xác định số lượng thông điệp SEI được lồng-định tỷ lệ được. Giá trị của `nesting_num_seis_minus1` sẽ nằm trong khoảng từ zero đến sáu mươi ba, bao gồm cả hai giá trị này. `nesting_zero_bit` sẽ được thiết lập bằng zero.

FIG.8 là sơ đồ giản lược của thiết bị lập mã video 800 làm ví dụ. Thiết bị lập mã video 800 thích hợp để thực hiện các ví dụ/phương án được mô tả như được mô tả ở đây. Thiết bị lập mã video 800 bao gồm các cổng xuôi dòng 820, các cổng ngược dòng 850, và/hoặc các bộ phận thu phát (Tx/Rx) 810, bao gồm các bộ phát và/hoặc bộ thu để truyền thông dữ liệu ngược dòng và/hoặc xuôi dòng qua mạng. Thiết bị lập mã video 800 còn bao gồm bộ xử lý 830 bao gồm bộ phận logic và/hoặc bộ xử lý trung tâm (CPU) để xử lý dữ liệu và bộ nhớ 832 để lưu trữ dữ liệu. Thiết bị lập mã video 800 cũng có thể bao gồm các thành phần điện, quang-thành-điện (OE), các thành phần điện-thành-quang (EO), và/hoặc các thành phần truyền thông không dây được nối với các cổng ngược dòng 850 và/hoặc các cổng xuôi dòng 820 để truyền thông dữ liệu thông qua các mạng

truyền thông điện, quang hoặc không dây. Thiết bị lập mã video 800 cũng có thể bao gồm các thiết bị đầu vào và/hoặc đầu ra (I/O) 860 để truyền thông dữ liệu đến và từ người dùng. Các thiết bị I/O 860 có thể bao gồm các thiết bị đầu ra như màn hình để hiển thị dữ liệu video, loa để xuất ra dữ liệu âm thanh, v.v.. Các thiết bị I/O 860 cũng có thể bao gồm các thiết bị đầu vào, như bàn phím, chuột, bi xoay, v.v., và/hoặc các giao diện tương ứng tương tác với các thiết bị đầu ra này.

Bộ xử lý 830 được thực hiện bởi phần cứng và phần mềm. Bộ xử lý 830 có thể được thực hiện dưới dạng một hoặc nhiều chip CPU, nhiều lõi (ví dụ bộ xử lý đa lõi), mảng cổng lập trình được theo trường (field-programmable gate array, FPGA), mạch tích hợp chuyên dụng (application specific integrated circuit, ASIC), và bộ xử lý tín hiệu số (digital signal processor, DSP). Bộ xử lý 830 truyền thông với các cổng xuôi dòng 820, Tx/Rx 810, các cổng ngược dòng 850, và bộ nhớ 832. Bộ xử lý 830 bao gồm môđun lập mã 814. Môđun lập mã 814 thực hiện các phương án được bộc lộ được mô tả ở trên, như các phương pháp 100, 900, và 1000, có thể sử dụng chuỗi video đa lớp 600 và/hoặc dòng bit 700. Môđun lập mã 814 cũng có thể thực hiện phương pháp/cơ chế bất kỳ khác được mô tả ở trên. Ngoài ra, môđun lập mã 814 cũng có thể thực hiện hệ thống lập-giải mã 200, bộ mã hóa 300, bộ giải mã 400, và/hoặc HRD 500. Ví dụ, môđun lập mã 814 có thể sử dụng tín hiệu và/hoặc đọc các tham số khác nhau như được mô tả ở trên. Ngoài ra, môđun lập mã cũng có thể được sử dụng để mã hóa và/hoặc giải mã chuỗi video dựa vào các tham số này. Do đó, việc báo hiệu các thay đổi được mô tả ở đây có thể làm tăng hiệu quả và/hoặc tránh các lỗi trong môđun lập mã 814. Do đó, môđun lập mã 814 có thể được tạo cấu hình để thực hiện các cơ chế để giải quyết một hoặc nhiều vấn đề được trình bày ở trên. Do đó, môđun lập mã 814 khiến thiết bị lập mã video 800 cung cấp chức năng bổ sung và/hoặc hiệu quả lập mã khi lập mã dữ liệu video. Do đó, môđun lập mã 814 nâng cao chức năng của thiết bị lập mã video 800 cũng như giải quyết được các vấn đề riêng cho kỹ thuật lập mã video. Ngoài ra, môđun lập mã 814 còn ảnh hưởng đến sự biến đổi của thiết bị lập mã video 800 sang một trạng thái khác. Theo cách khác, môđun lập mã 814 có thể được thực hiện dưới dạng các lệnh được lưu trữ trong bộ nhớ 832 và được thực thi bởi bộ xử lý 830 (ví dụ dưới dạng sản phẩm chương trình máy tính được lưu trữ trên phương tiện không tạm thời).

Bộ nhớ 832 bao gồm một hoặc nhiều kiểu bộ nhớ như đĩa, ổ băng, ổ đĩa thể rắn, bộ nhớ chỉ đọc (read-only memory, ROM), bộ nhớ truy cập ngẫu nhiên (random access memory, RAM), bộ nhớ chớp, nội dung khả lập địa chỉ tam phân (ternary content-addressable memory, TCAM), bộ nhớ truy cập ngẫu nhiên tĩnh (static random-access memory, SRAM), v.v.. Bộ nhớ 832 có thể được sử dụng làm thiết bị lưu trữ dữ liệu tràn, để lưu trữ các chương trình khi các chương trình này được chọn để thực thi, và lưu trữ các lệnh và dữ liệu được đọc trong quá trình thực thi chương trình.

FIG.9 là lưu đồ của phương pháp 900 làm ví dụ để mã hóa chuỗi video thành dòng bit, như dòng bit 700, bằng cách ràng buộc các TemporalId cho các thông điệp SEI trong dòng bit. Phương pháp 900 có thể được sử dụng bởi bộ mã hóa, như hệ thống lập-giải mã 200, bộ mã hóa 300, và/hoặc thiết bị lập mã video 800 khi thực hiện phương pháp 100. Ngoài ra, phương pháp 900 có thể hoạt động trên HRD 500 và do đó có thể thực hiện phép kiểm tra sự tương thích trên chuỗi video đa lớp 600.

Phương pháp 900 có thể bắt đầu khi bộ mã hóa nhận chuỗi video và xác định để mã hóa chuỗi video này thành dòng bit đa lớp, ví dụ dựa vào đầu vào người dùng. Ở bước 901, bộ mã hóa mã hóa hình ảnh được lập mã trong một hoặc nhiều đơn vị VCL NAL trong dòng bit. Ví dụ, hình ảnh được lập mã có thể có trong AU trong một lớp. Ngoài ra, bộ mã hóa cũng có thể mã hóa một hoặc nhiều lớp bao gồm hình ảnh được lập mã thành dòng bit đa lớp. Một lớp có thể bao gồm tập hợp các đơn vị VCL NAL có cùng ID lớp và các đơn vị non-VCL NAL có liên quan. Ví dụ, tập hợp các đơn vị VCL NAL là một phần của một lớp khi tập hợp các đơn vị VCL NAL đều có giá trị cụ thể là `nuh_layer_id`. Một lớp có thể bao gồm tập hợp các đơn vị VCL NAL chứa dữ liệu video của các hình ảnh được mã hóa cũng như các tập hợp tham số bất kỳ được sử dụng để lập mã các hình ảnh này. Một hoặc nhiều lớp có thể là các lớp đầu ra. Các lớp không phải lớp đầu ra được mã hóa để hỗ trợ việc tái dựng (các) lớp đầu ra, mà việc hỗ trợ các lớp này không dự định dành cho đầu ra ở bộ giải mã. Bằng cách này, bộ mã hóa có thể mã hóa các tổ hợp khác nhau của các lớp để truyền đến bộ giải mã theo yêu cầu. Lớp này có thể được truyền theo mong muốn để cho phép bộ giải mã thu nhận các biểu diễn khác nhau của chuỗi video phụ thuộc vào các điều kiện mạng, các khả năng phần cứng, và/hoặc các thiết lập người dùng.

Ở bước 903, bộ mã hóa có thể mã hóa một hoặc nhiều đơn vị non-VCL NAL thành dòng bit. Ví dụ, lớp và/hoặc tập hợp các lớp còn bao gồm các đơn vị non-VCL NAL khác nhau. Các đơn vị non-VCL NAL được liên kết với tập hợp các đơn vị VCL NAL đều có giá trị cụ thể là `nuh_layer_id`. Cụ thể là, đơn vị non-VCL NAL được mã hóa sao cho `TemporalId` cho đơn vị non-VCL NAL được ràng buộc bằng `TemporalId` của AU chứa đơn vị non-VCL NAL khi `nal_unit_type` của non-VCL NAL biểu thị thông điệp SEI có trong non-VCL NAL. Nói cách khác, thông điệp SEI có thể có trong cùng AU dưới dạng hình ảnh mà SEI áp dụng. Do đó, `TemporalId` của thông điệp SEI chứa trong đơn vị non-VCL NAL được ràng buộc bằng `TemporalId` của AU chứa thông điệp SEI/ đơn vị non-VCL NAL. Trong một số ví dụ, thông điệp SEI là thông điệp SEI tiền tố, và do đó `nal_unit_type` của non-VCL NAL bằng `PREFIX_SEI_NUT`. Trong một số ví dụ, thông điệp SEI là thông điệp SEI hậu tố, và do đó `nal_unit_type` của non-VCL NAL bằng `SUFFIX_SEI_NUT`.

`TemporalId` cho đơn vị non-VCL NAL có thể được xác định bởi phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị non-VCL NAL. Tương tự, `TemporalId` cho AU có thể được xác định bởi phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị VCL NAL chứa lát của hình ảnh được lập mã trong AU. `TemporalId` của các đơn vị VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU. Do đó, phần tử cú pháp `nuh_temporal_id_plus1` của các đơn vị VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU. Do đó, giá trị của phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị non-VCL NAL chứa thông điệp SEI giống như giá trị của phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị VCL NAL bất kỳ trong cùng AU dưới dạng thông điệp SEI. `TemporalId` cho đơn vị non-VCL NAL được suy ra như sau:

$$\text{TemporalId} = \text{nuh_temporal_id_plus1} - 1.$$

Ngoài ra, giá trị của `nuh_temporal_id_plus1` cho đơn vị non-VCL NAL và đơn vị VCL NAL trong AU có thể không được thiết lập bằng zero do điều này sẽ dẫn đến giá trị âm của `TemporalId`. Ví dụ, các ràng buộc và/hoặc yêu cầu có trước bảo đảm rằng dòng bit tương thích với chuẩn VVC hoặc một số chuẩn khác, được cải biến như được biểu thị ở đây. Tuy nhiên, bộ mã hóa cũng có thể có khả năng hoạt động theo các chế độ khác trong đó nó không bị ràng buộc như vậy, như khi hoạt động dưới chuẩn khác hoặc

phiên bản khác của cùng chuẩn.

Ở bước 905, bộ mã hóa sử dụng HRD để thực hiện tập hợp các phép kiểm tra sự tương thích dòng bit trên dòng bit này dựa vào thông điệp SEI. Tập hợp này có thể bao gồm một hoặc nhiều phép kiểm tra sự tương thích. Ví dụ, HRD có thể sử dụng các TemporalId và/hoặc các giá trị `nuh_temporal_id_plus1` để phối hợp các thông điệp SEI với các hình ảnh. Do đó, HRD có thể sử dụng các tham số từ thông điệp SEI để thực hiện một hoặc nhiều phép kiểm tra sự tương thích trên hình ảnh được lập mã trong cùng AU dưới dạng thông điệp SEI. Sau đó, bộ mã hóa có thể lưu trữ dòng bit để truyền thông về phía bộ giải mã ở bước 907. Bộ mã hóa cũng có thể truyền dòng bit về phía bộ giải mã theo mong muốn.

FIG.10 là lưu đồ của phương pháp 1000 làm ví dụ để giải mã chuỗi video từ một dòng bit, như dòng bit 700, trong đó các TemporalId cho các thông điệp SEI trong dòng bit được ràng buộc. Phương pháp 1000 có thể được sử dụng bởi bộ giải mã, như hệ thống lập-giải mã 200, bộ giải mã 400, và/hoặc thiết bị lập mã video 800 khi thực hiện phương pháp 100. Ngoài ra, phương pháp 1000 có thể được sử dụng trên chuỗi video đa lớp 600 đã được kiểm tra về sự tương thích bởi HRD, như HRD 500.

Phương pháp 1000 có thể bắt đầu khi bộ giải mã bắt đầu nhận dòng bit của dữ liệu được lập mã biểu diễn chuỗi video đa lớp, ví dụ dưới dạng kết quả của phương pháp 900 và/hoặc để đáp lại yêu cầu bởi bộ giải mã. Ở bước 1001, bộ giải mã nhận dòng bit dòng bit bao gồm hình ảnh được lập mã trong một hoặc nhiều đơn vị VCL NAL và đơn vị non-VCL NAL. Ví dụ, hình ảnh được lập mã có thể có trong AU. Ngoài ra, dòng bit có thể bao gồm một hoặc nhiều lớp bao gồm hình ảnh được lập mã. Một lớp có thể bao gồm tập hợp các đơn vị VCL NAL có cùng ID lớp và các đơn vị non-VCL NAL có liên quan. Ví dụ, tập hợp các đơn vị VCL NAL là một phần của lớp khi tập hợp các đơn vị VCL NAL đều có giá trị cụ thể là `nuh_layer_id`. Một lớp có thể bao gồm tập hợp các đơn vị VCL NAL chứa dữ liệu video của các hình ảnh được lập mã cũng như các tập hợp tham số bất kỳ được sử dụng để lập mã các hình ảnh này. Một hoặc nhiều lớp này có thể là các lớp đầu ra. Các lớp này không phải là lớp đầu ra được mã hóa để hỗ trợ việc tái dựng (các) lớp đầu ra, mà việc hỗ trợ các lớp này không dự định dành cho đầu ra. Bằng cách này, bộ giải mã có thể thu nhận các biểu diễn khác nhau của chuỗi video tùy thuộc vào các điều kiện mạng, các khả năng phần cứng, và/hoặc các thiết lập người

dùng. Lớp này cũng bao gồm các đơn vị non-VCL NAL khác nhau. Các đơn vị non-VCL NAL được liên kết với tập hợp các đơn vị VCL NAL đều có giá trị cụ thể là `nuh_layer_id`.

Cụ thể là, đơn vị non-VCL NAL được lập mã trong dòng bit sao cho TemporalId cho đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL khi `nal_unit_type` của non-VCL NAL biểu thị thông điệp SEI có trong non-VCL NAL. Nói cách khác, thông điệp SEI có thể có trong cùng AU dưới dạng hình ảnh mà SEI áp dụng. Do đó, TemporalId của thông điệp SEI chứa trong đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa thông điệp SEI/ đơn vị non-VCL NAL. Trong một số ví dụ, thông điệp SEI là thông điệp SEI tiên tố, và do đó `nal_unit_type` của non-VCL NAL bằng PREFIX_SEI_NUT. Trong một số ví dụ, thông điệp SEI là thông điệp SEI hậu tố, và do đó `nal_unit_type` của non-VCL NAL bằng SUFFIX_SEI_NUT.

TemporalId cho đơn vị non-VCL NAL có thể được xác định bởi phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị non-VCL NAL. Do đó, bộ giải mã có thể suy ra TemporalId cho đơn vị non-VCL NAL dựa vào phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị non-VCL NAL ở bước 1002. Tương tự, TemporalId cho AU có thể được xác định bởi phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị VCL NAL chứa lát của hình ảnh được lập mã trong AU. TemporalId của các đơn vị VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU. Do đó, phần tử cú pháp `nuh_temporal_id_plus1` của các đơn vị VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU. Do đó, giá trị của phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị non-VCL NAL chứa thông điệp SEI giống như giá trị của phần tử cú pháp `nuh_temporal_id_plus1` trong đơn vị VCL NAL bất kỳ trong cùng AU dưới dạng thông điệp SEI. TemporalId cho đơn vị non-VCL NAL được suy ra như sau:

$$\text{TemporalId} = \text{nuh_temporal_id_plus1} - 1.$$

Ngoài ra, giá trị của `nuh_temporal_id_plus1` cho đơn vị non-VCL NAL và đơn vị VCL NAL trong AU có thể không được thiết lập bằng zero do điều này sẽ dẫn đến giá trị âm của TemporalId.

Theo một phương án, bộ giải mã video mong muốn TemporalId cho đơn vị non-VCL NAL bằng TemporalId của AU chứa đơn vị non-VCL NAL khi nal_unit_type của non-VCL NAL là thông điệp SEI như được mô tả ở trên dựa vào chuẩn VVC hoặc một số chuẩn khác. Tuy nhiên, nếu bộ giải mã xác định rằng điều kiện này không đúng, bộ giải mã có thể phát hiện lỗi, báo hiệu lỗi, yêu cầu rằng dòng bit sửa đổi (hoặc một phần của nó) được gửi lại, hoặc thực hiện một số biện pháp hiệu chỉnh để bảo đảm rằng dòng bit thích hợp được nhận.

Ở bước 1003, bộ giải mã có thể giải mã hình ảnh được lập mã từ các đơn vị VCL NAL để tạo ra hình ảnh được giải mã. Ví dụ, bộ giải mã có thể sử dụng các TemporalId và/hoặc các giá trị nuh_temporal_id_plus1 để phối hợp các thông điệp SEI với các hình ảnh. Sau đó, bộ giải mã có thể sử dụng các thông điệp SEI theo mong muốn khi giải mã hình ảnh được lập mã. Ở bước 1005, bộ giải mã có thể chuyển tiếp hình ảnh được giải mã để hiển thị dưới dạng một phần của chuỗi video được giải mã.

FIG.11 là sơ đồ giản lược của hệ thống 1100 làm ví dụ để lập mã chuỗi video sử dụng dòng bit trong đó các TemporalId cho các thông điệp SEI trong dòng bit được ràng buộc. Hệ thống 1100 có thể được thực hiện bởi bộ mã hóa và bộ giải mã như hệ thống lập-giải mã 200, bộ mã hóa 300, bộ giải mã 400, và/hoặc thiết bị lập mã video 800. Ngoài ra, hệ thống 1100 cũng có thể sử dụng HRD 500 để thực hiện phép kiểm tra sự tương thích trên chuỗi video đa lớp 600 và/hoặc dòng bit 700. Ngoài ra, hệ thống 1100 cũng có thể được sử dụng khi thực hiện phương pháp 100, 900, và/hoặc 1000.

Hệ thống 1100 bao gồm bộ mã hóa video 1102. Bộ mã hóa video 1102 bao gồm môđun mã hóa 1103 để mã hóa hình ảnh được lập mã trong một hoặc nhiều đơn vị VCL NAL trong dòng bit. Môđun mã hóa 1103 còn để mã hóa thành dòng bit đơn vị non-VCL NAL sao cho TemporalId cho đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL khi nal_unit_type của non-VCL NAL là thông điệp SEI. Bộ mã hóa video 1102 còn bao gồm môđun HRD 1105 để thực hiện tập hợp các phép kiểm tra sự tương thích dòng bit trên dòng bit dựa vào thông điệp SEI. Bộ mã hóa video 1102 còn bao gồm môđun lưu trữ 1106 để lưu trữ dòng bit để truyền thông về phía bộ giải mã. Bộ mã hóa video 1102 còn bao gồm môđun phát 1107 để phát dòng bit về phía bộ giải mã video 1110. Bộ mã hóa video 1102 có thể còn được tạo cấu hình để thực hiện bước bất kỳ trong số các bước của phương pháp 900.

Hệ thống 1100 còn bao gồm bộ giải mã video 1110. Bộ giải mã video 1110 bao gồm môđun thu 1111 để thu dòng bit bao gồm hình ảnh được lập mã trong một hoặc nhiều đơn vị VCL NAL và đơn vị non-VCL NAL, trong đó TemporalId cho đơn vị non-VCL NAL được ràng buộc bằng TemporalId của AU chứa đơn vị non-VCL NAL khi nal_unit_type của non-VCL NAL là thông điệp SEI. Bộ giải mã video 1110 còn bao gồm môđun giải mã 1113 để giải mã hình ảnh được lập mã từ các đơn vị VCL NAL để tạo ra hình ảnh được giải mã. Bộ giải mã video 1110 còn bao gồm môđun chuyển tiếp 1115 để chuyển tiếp hình ảnh được giải mã để hiển thị dưới dạng một phần của chuỗi video được giải mã. Bộ giải mã video 1110 cũng có thể được tạo cấu hình để thực hiện bước bất kỳ trong số các bước của phương pháp 1000.

Thành phần thứ nhất được nối trực tiếp với thành phần thứ hai khi không có các thành phần xen giữa, ngoại trừ đường dẫn, vết hoặc phương tiện khác giữa thành phần thứ nhất và thành phần thứ hai. Thành phần thứ nhất được nối gián tiếp với thành phần thứ hai khi có các thành phần xen giữa không phải đường dẫn, vết, hoặc phương tiện khác giữa thành phần thứ nhất và thành phần thứ hai. Thuật ngữ “được nối” và các biến thể của nó bao gồm cả nối trực tiếp lẫn nối gián tiếp. Việc sử dụng thuật ngữ “khoảng” nghĩa là phạm vi bao gồm $\pm 10\%$ của số tiếp theo trừ khi được quy định khác.

Cũng cần hiểu rằng các bước của các phương pháp làm ví dụ nêu trên không nhất thiết yêu cầu được thực hiện theo thứ tự được mô tả, và thứ tự của các bước của các phương pháp này nên được hiểu chỉ là ví dụ. Tương tự, các bước bổ sung có thể có trong các phương pháp này, và các bước nhất định có thể được lược bỏ hoặc kết hợp, trong các phương pháp theo các phương án khác nhau của sáng chế.

Mặc dù một số phương án đã được đề xuất theo sáng chế, nhưng cần hiểu rằng các hệ thống và các phương pháp được mô tả có thể được thực hiện theo nhiều dạng cụ thể khác mà không vượt ra khỏi phạm vi của sáng chế. Các ví dụ này cần được xem là minh họa và không nhằm giới hạn, và dự định không bị giới hạn ở nội dung chi tiết nêu trên. Ví dụ, các phần tử hoặc các thành phần khác nhau có thể được kết hợp hoặc được tích hợp trong hệ thống khác hoặc các dấu hiệu nhất định có thể được lược bỏ, hoặc không được thực hiện.

Ngoài ra, các kỹ thuật, hệ thống, hệ thống con, và các phương pháp được mô tả và được minh họa theo các phương án khác nhau dưới dạng rời rạc hoặc tách biệt

có thể được kết hợp hoặc được tích hợp với các hệ thống, thành phần, kỹ thuật, hoặc các phương pháp khác mà không vượt ra khỏi phạm vi của sáng chế. Các ví dụ khác về các thay đổi, thay thế và sửa đổi xác định được bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này và có thể được thực hiện mà không vượt ra khỏi phạm vi được bộc lộ ở đây.

YÊU CẦU BẢO HỘ

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

nhận, bởi bộ thu của bộ giải mã, dòng bit bao gồm hình ảnh được lập mã trong một hoặc nhiều đơn vị lớp trừu tượng hóa mạng loại lớp lập mã video (video coding layer (VCL) network abstraction layer (NAL) unit) và đơn vị non-VCL NAL, trong đó mã định danh theo thời gian (TemporalId) cho đơn vị non-VCL NAL bằng TemporalId của đơn vị truy cập (access unit, AU) chứa đơn vị non-VCL NAL khi kiểu đơn vị NAL (nal_unit_type) của đơn vị non-VCL NAL bằng kiểu đơn vị NAL thông tin tăng cường bổ sung (supplemental enhancement information (SEI) NAL unit type) tiền tố (PREFIX_SEI_NUT) hoặc kiểu đơn vị SEI NAL hậu tố (SUFFIX_SEI_NUT), trong đó khi thông điệp SEI lồng định tỷ lệ được bao gồm kiểu trọng tải được thiết lập thành một trăm ba mươi ba, nal_unit_type của đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được bằng PREFIX_SEI_NUT hoặc SUFFIX_SEI_NUT;

suy ra, bởi bộ xử lý của bộ giải mã, TemporalId cho đơn vị non-VCL NAL dựa vào phần tử cú pháp mã định danh theo thời gian phần đầu đơn vị NAL cộng một (nuh_temporal_id_plus1) trong đơn vị non-VCL NAL; và

giải mã, bởi bộ xử lý của bộ giải mã, hình ảnh được lập mã từ các đơn vị VCL NAL để tạo ra hình ảnh được giải mã.

2. Phương pháp theo điểm 1, trong đó nal_unit_type của non-VCL NAL bằng PREFIX_SEI_NUT.

3. Phương pháp theo điểm 1, trong đó nal_unit_type của non-VCL NAL bằng SUFFIX_SEI_NUT.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó hình ảnh được lập mã được giải mã từ các đơn vị VCL NAL dựa vào thông điệp SEI trong đơn vị non-VCL NAL.

5. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 4, trong đó phương pháp này còn bao gồm bước suy ra TemporalId cho đơn vị non-VCL NAL như sau:

$$\text{TemporalId} = \text{nuh_temporal_id_plus1} - 1.$$

6. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó giá trị của nuh_temporal_id_plus1 không bằng zero.

7. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong đó TemporalId của các đơn vị VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU.

8. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 7, trong đó phương pháp này còn bao gồm bước:

nhận, bởi bộ giải mã, dòng bit thứ hai bao gồm một hoặc nhiều đơn vị VCL NAL thứ hai và một đơn vị non-VCL NAL thứ hai, trong đó TemporalId cho đơn vị non-VCL NAL thứ hai không bằng TemporalId của AU thứ hai chứa đơn vị non-VCL NAL thứ hai khi nal_unit_type của non-VCL NAL thứ hai là thông điệp SEI; và

đáp lại bước nhận nêu trên, thực hiện một biện pháp hiệu chỉnh để bảo đảm rằng một dòng bit thích hợp tương ứng với dòng bit thứ hai được nhận trước khi giải mã hình ảnh được lập mã từ các đơn vị VCL NAL thứ hai.

9. Phương tiện lưu trữ đọc được bởi máy tính có chương trình được ghi trên đó; trong đó chương trình này khiến cho máy tính thực thi phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 8.

10. Thiết bị giải mã bao gồm:

bộ thu được tạo cấu hình để nhận dòng bit bao gồm hình ảnh được lập mã trong một hoặc nhiều đơn vị lớp trừu tượng hóa mạng loại lớp lập mã video (video coding layer (VCL) network abstraction layer (NAL) unit) và đơn vị non-VCL NAL, trong đó mã định danh theo thời gian (TemporalId) cho đơn vị non-VCL NAL bằng TemporalId của đơn vị truy cập (access unit, AU) chứa đơn vị non-VCL NAL khi kiểu đơn vị NAL (nal_unit_type) của đơn vị non-VCL NAL là thông điệp thông tin tăng cường bổ sung (supplemental enhancement information, SEI), trong đó khi thông điệp SEI lồng định tỷ lệ được bao gồm kiểu trọng tải được thiết lập thành một trăm ba mươi ba, nal_unit_type của đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được bằng PREFIX_SEI_NUT hoặc SUFFIX_SEI_NUT;

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

bộ xử lý được nối với bộ nhớ, bộ xử lý này được tạo cấu hình để thực thi các lệnh để khiến cho thiết bị giải mã: thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 8.

11. Phương pháp mã hóa được thực hiện bởi bộ mã hóa, phương pháp này bao gồm các bước:

mã hóa hình ảnh được lập mã trong một hoặc nhiều đơn vị lớp trừu tượng hóa mạng loại lớp lập mã video (video coding layer (VCL) network abstraction layer (NAL) unit) trong dòng bit;

mã hóa thành dòng bit đơn vị non-VCL NAL;

trong đó mã định danh theo thời gian (TemporalId) cho đơn vị non-VCL NAL bằng TemporalId của đơn vị truy cập (access unit, AU) chứa đơn vị non-VCL NAL khi kiểu đơn vị NAL (nal_unit_type) của đơn vị non-VCL NAL bằng kiểu đơn vị NAL thông tin tăng cường bổ sung (supplemental enhancement information (SEI) NAL unit type) tiền tố (PREFIX_SEI_NUT) hoặc kiểu đơn vị SEI NAL hậu tố (SUFFIX_SEI_NUT), trong đó khi thông điệp SEI lồng định tỷ lệ được bao gồm kiểu trọng tải được thiết lập thành một trăm ba mươi ba, nal_unit_type của đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được bằng PREFIX_SEI_NUT hoặc SUFFIX_SEI_NUT;

trong đó TemporalId cho đơn vị non-VCL NAL được suy ra dựa vào phần tử cú pháp mã định danh theo thời gian phần đầu đơn vị NAL cộng một (nuh_temporal_id_plus1) trong đơn vị non-VCL NAL.

12. Phương pháp theo điểm 11, trong đó nal_unit_type của đơn vị non-VCL NAL bằng PREFIX_SEI_NUT.

13. Phương pháp theo điểm 11, trong đó nal_unit_type của đơn vị non-VCL NAL bằng SUFFIX_SEI_NUT.

14. Phương pháp theo điểm bất kỳ trong số các điểm từ 11 đến 13, trong đó TemporalId cho đơn vị non-VCL NAL được suy ra như sau:

$$\text{TemporalId} = \text{nuh_temporal_id_plus1} - 1.$$

15. Phương pháp theo điểm bất kỳ trong số các điểm từ 11 đến 14, trong đó giá trị của nuh_temporal_id_plus1 không bằng zero.

16. Phương pháp theo điểm bất kỳ trong số các điểm từ 11 đến 15, trong đó nuh_temporal_id_plus1 của các đơn vị VCL NAL được ràng buộc giống nhau đối với tất cả các đơn vị VCL NAL trong cùng AU.

17. Phương tiện đọc được bằng máy tính không tạm thời lưu trữ các lệnh máy tính mà khi được thực thi bởi ít nhất một bộ xử lý phần cứng, khiến ít nhất một bộ xử lý phần cứng thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 11 đến 16.

18. Thiết bị mã hóa bao gồm:

bộ nhớ và bộ xử lý, trong đó bộ nhớ được ghép nối với bộ xử lý, bộ nhớ lưu trữ các lệnh; và

bộ xử lý được tạo cấu hình để thực thi các lệnh để khiến thiết bị mã hóa: thực hiện phương pháp theo điểm bất kỳ trong số các điểm từ 11 đến 16.

19. Phương tiện lưu trữ không tạm thời bao gồm dòng bit, trong đó dòng bit này bao gồm: hình ảnh được lập mã trong một hoặc nhiều đơn vị lớp trừu tượng hóa mạng loại lớp lập mã video (video coding layer (VCL) network abstraction layer (NAL) unit) và đơn vị non-VCL NAL, trong đó mã định danh theo thời gian (TemporalId) cho đơn vị non-VCL NAL bằng TemporalId của đơn vị truy cập (access unit, AU) chứa đơn vị non-VCL NAL khi kiểu đơn vị NAL (nal_unit_type) của đơn vị non-VCL NAL bằng kiểu đơn vị NAL thông tin tăng cường bổ sung (supplemental enhancement information (SEI) NAL unit type) tiền tố (PREFIX_SEI_NUT) hoặc kiểu đơn vị SEI NAL hậu tố (SUFFIX_SEI_NUT), trong đó khi thông điệp SEI lồng định tỷ lệ được bao gồm kiểu trọng tải được thiết lập thành một trăm ba mươi ba, nal_unit_type của đơn vị SEI NAL chứa thông điệp SEI lồng định tỷ lệ được bằng PREFIX_SEI_NUT hoặc SUFFIX_SEI_NUT.

1/11

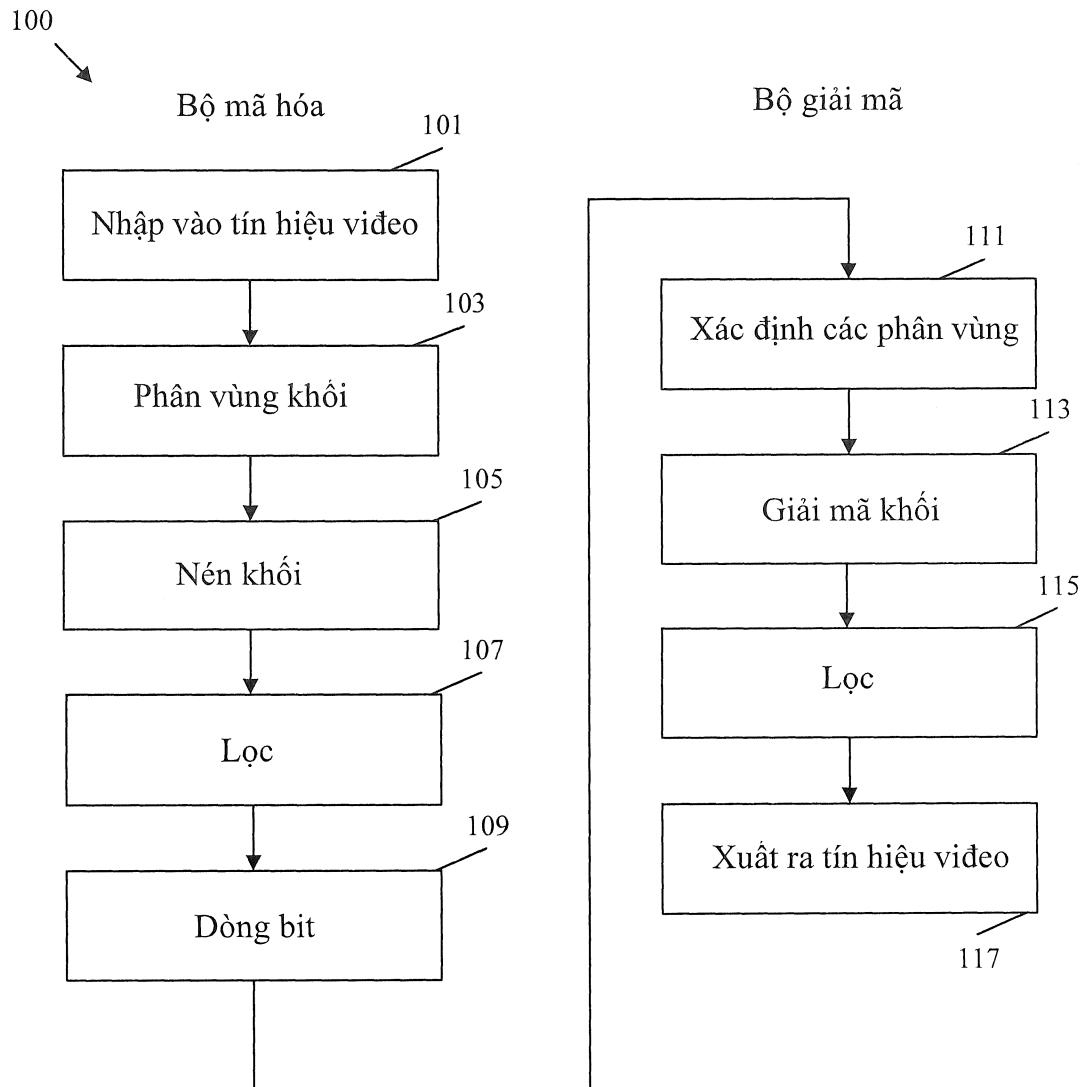


FIG. 1

2/11

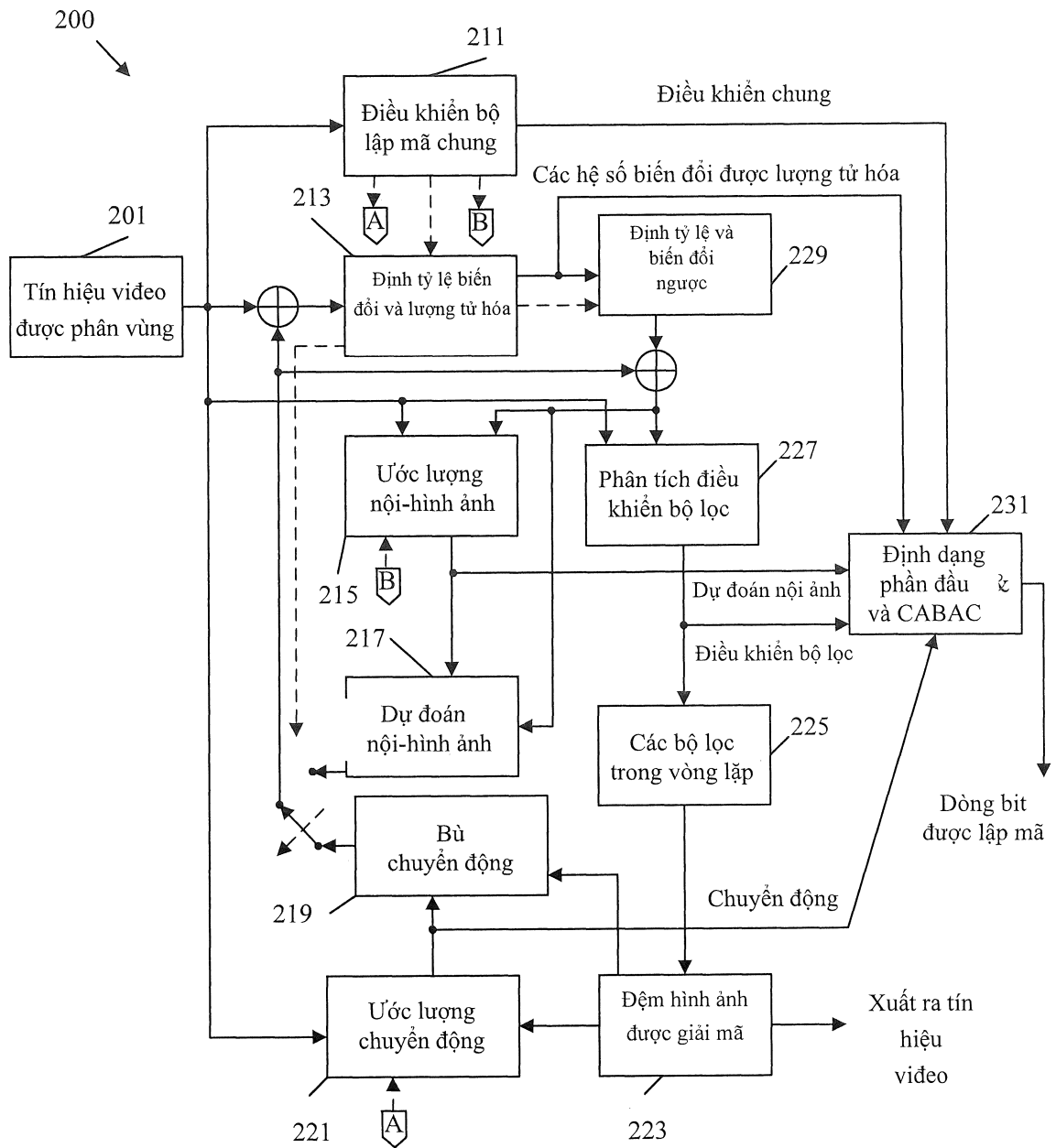


FIG. 2

3/11

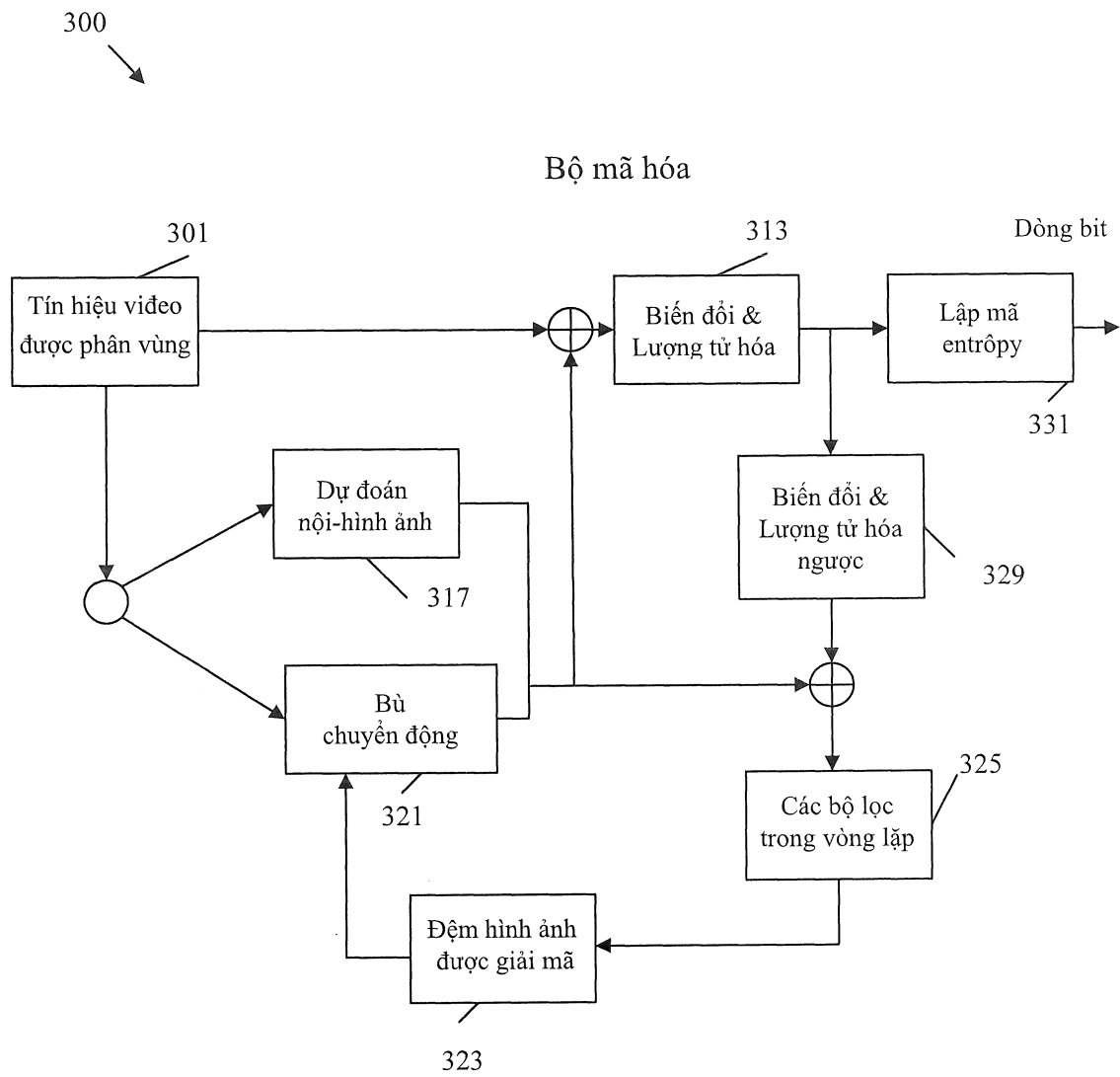


FIG. 3

4/11

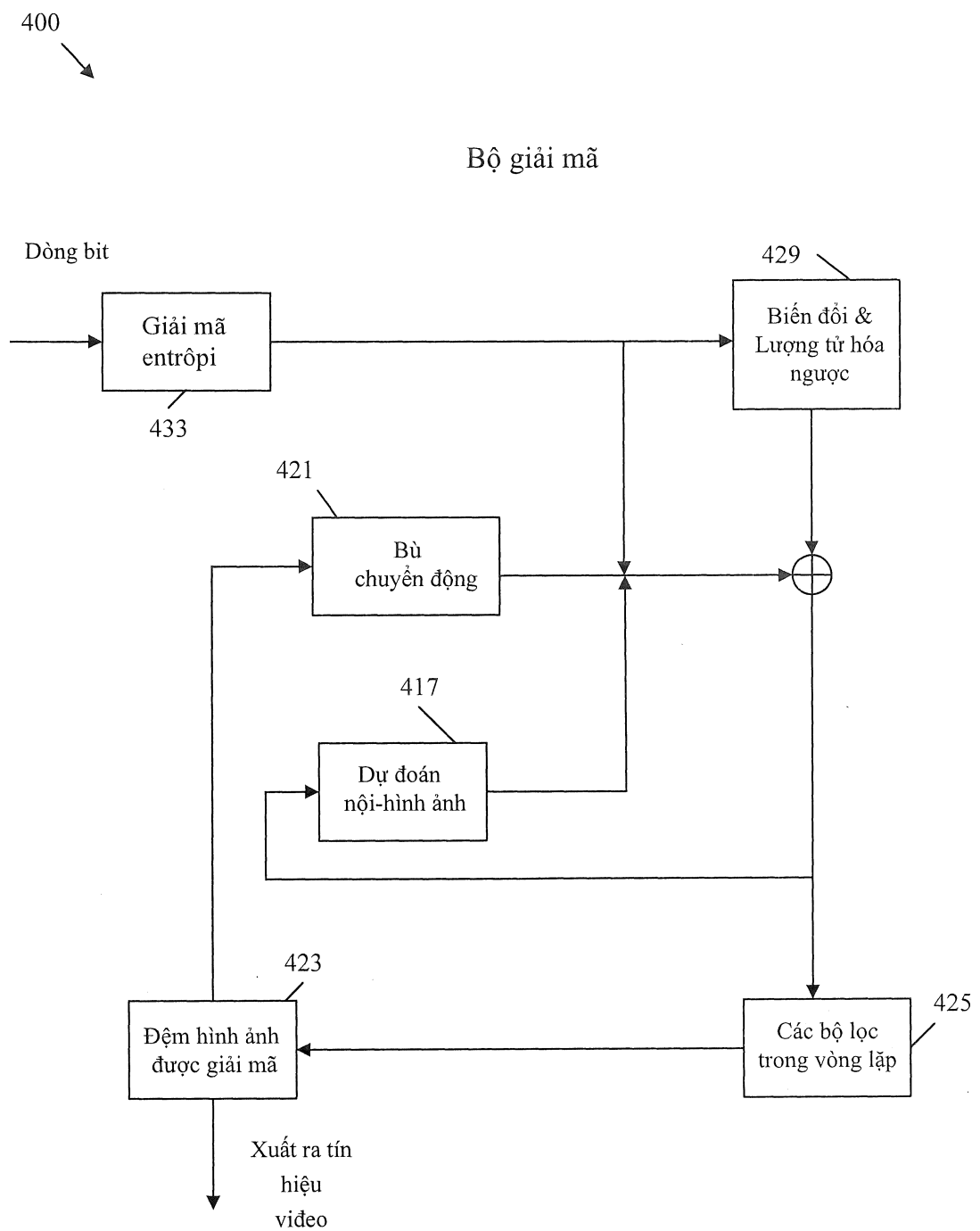


FIG. 4

5/11

Bộ giải mã tham chiếu giả định

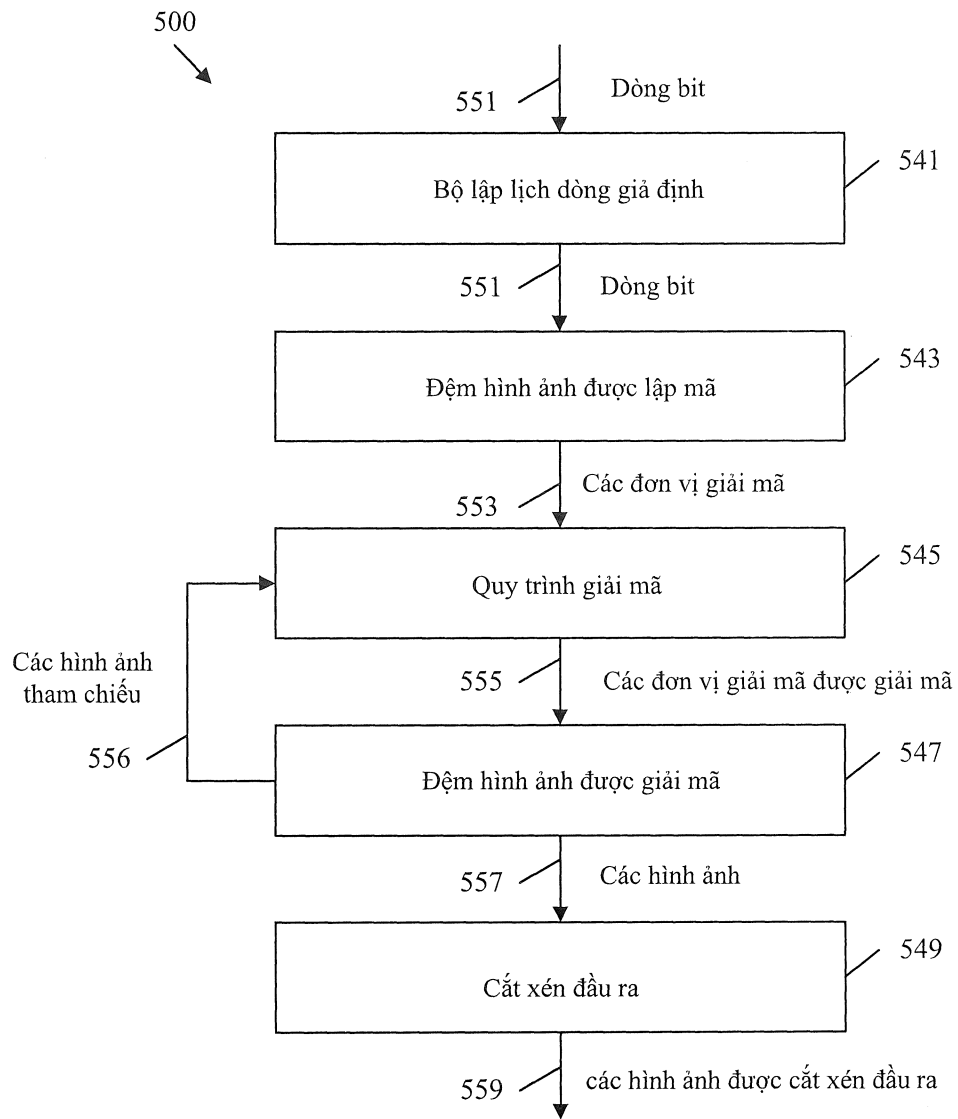


FIG. 5

6/11

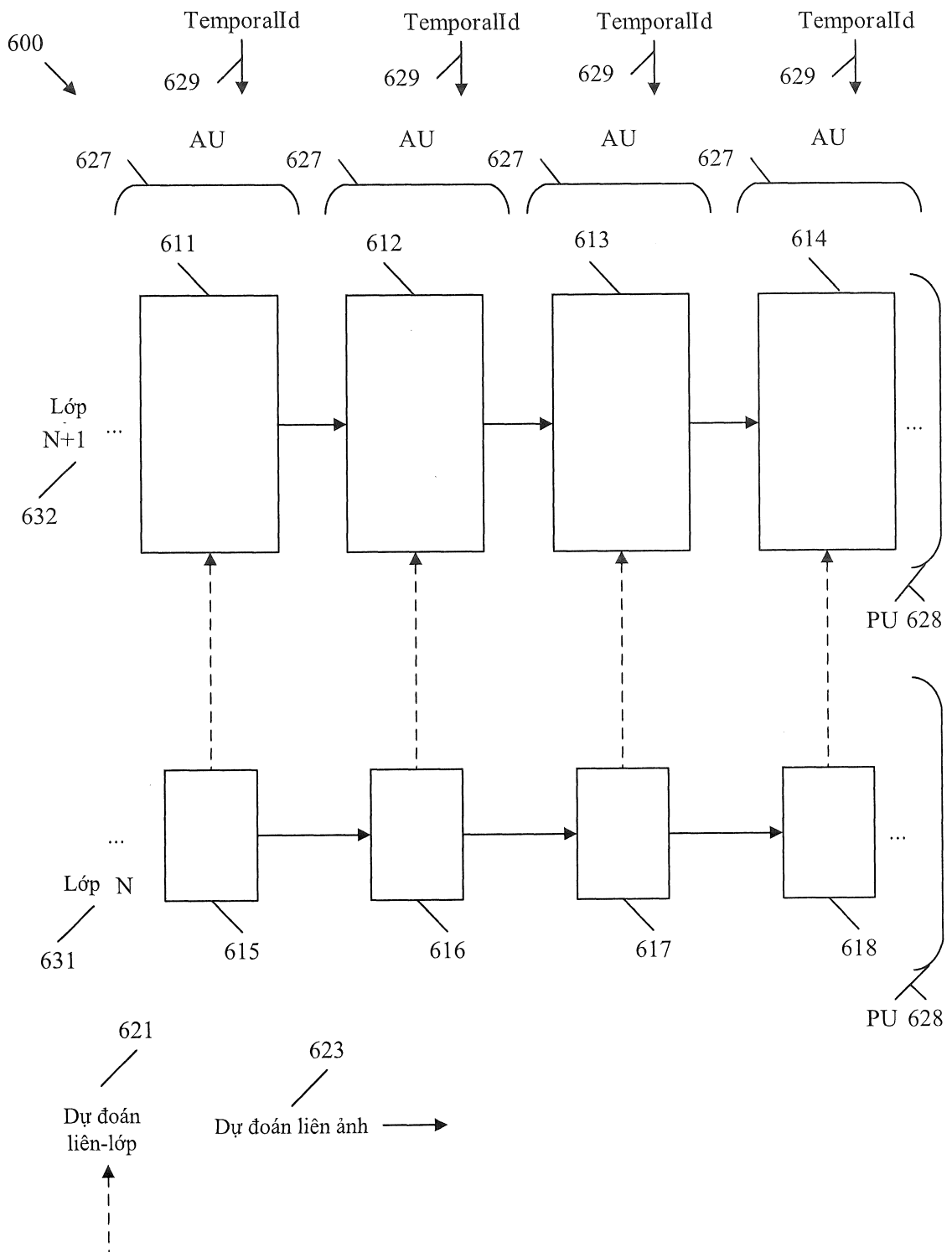


FIG. 6

7/11

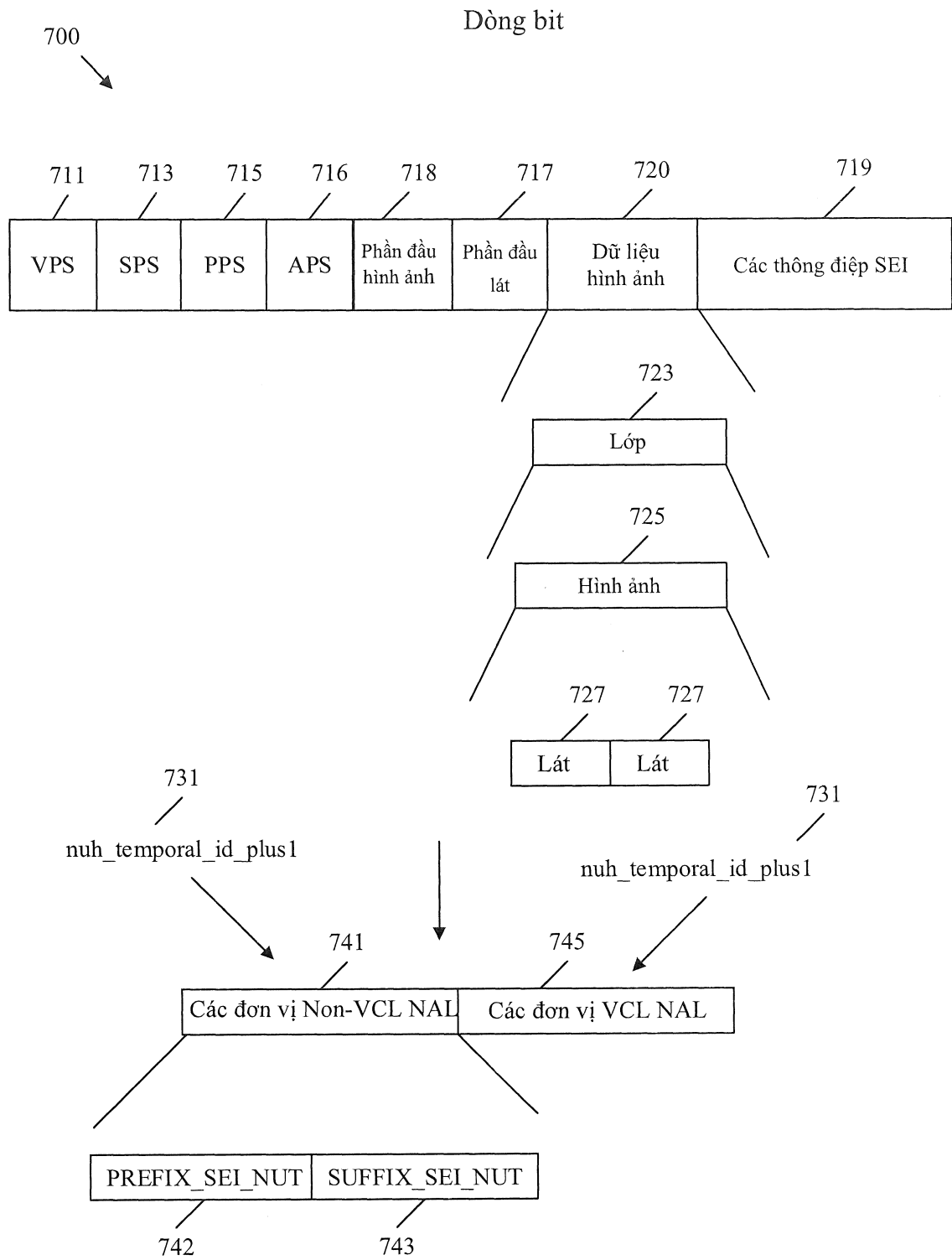


FIG. 7

8/11

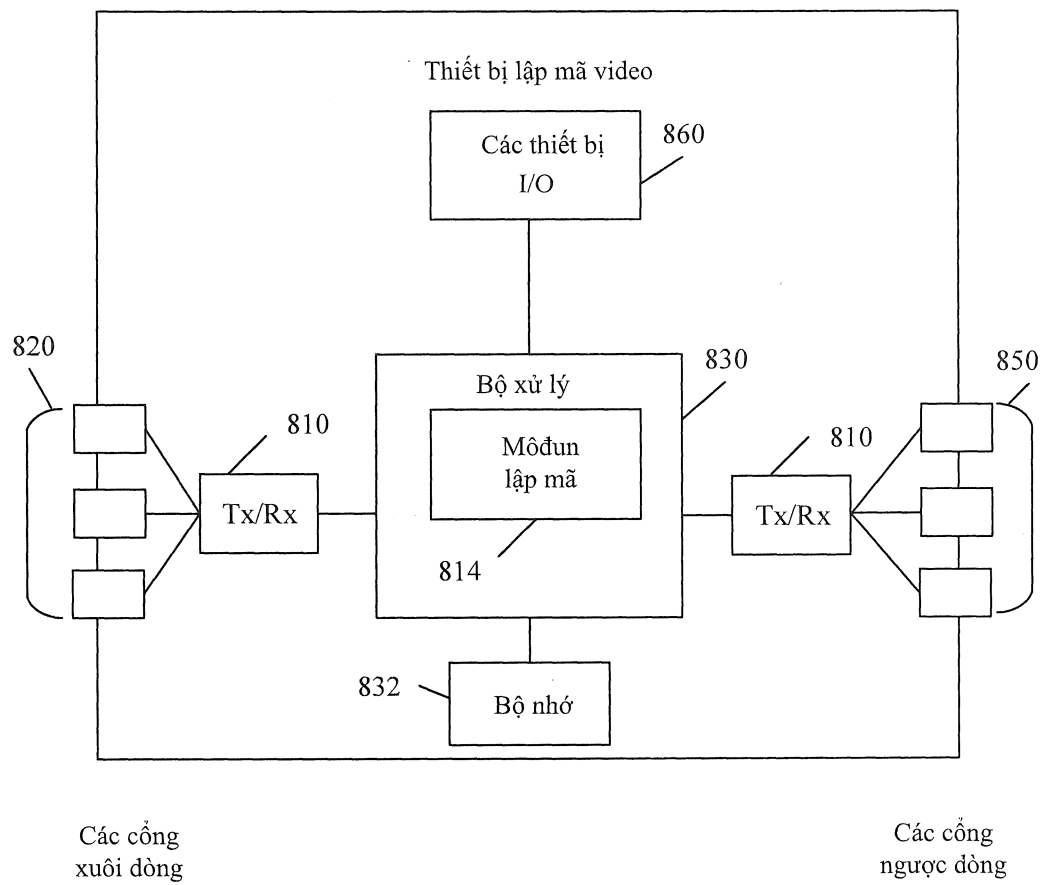
800
↓

FIG. 8

9/11

900

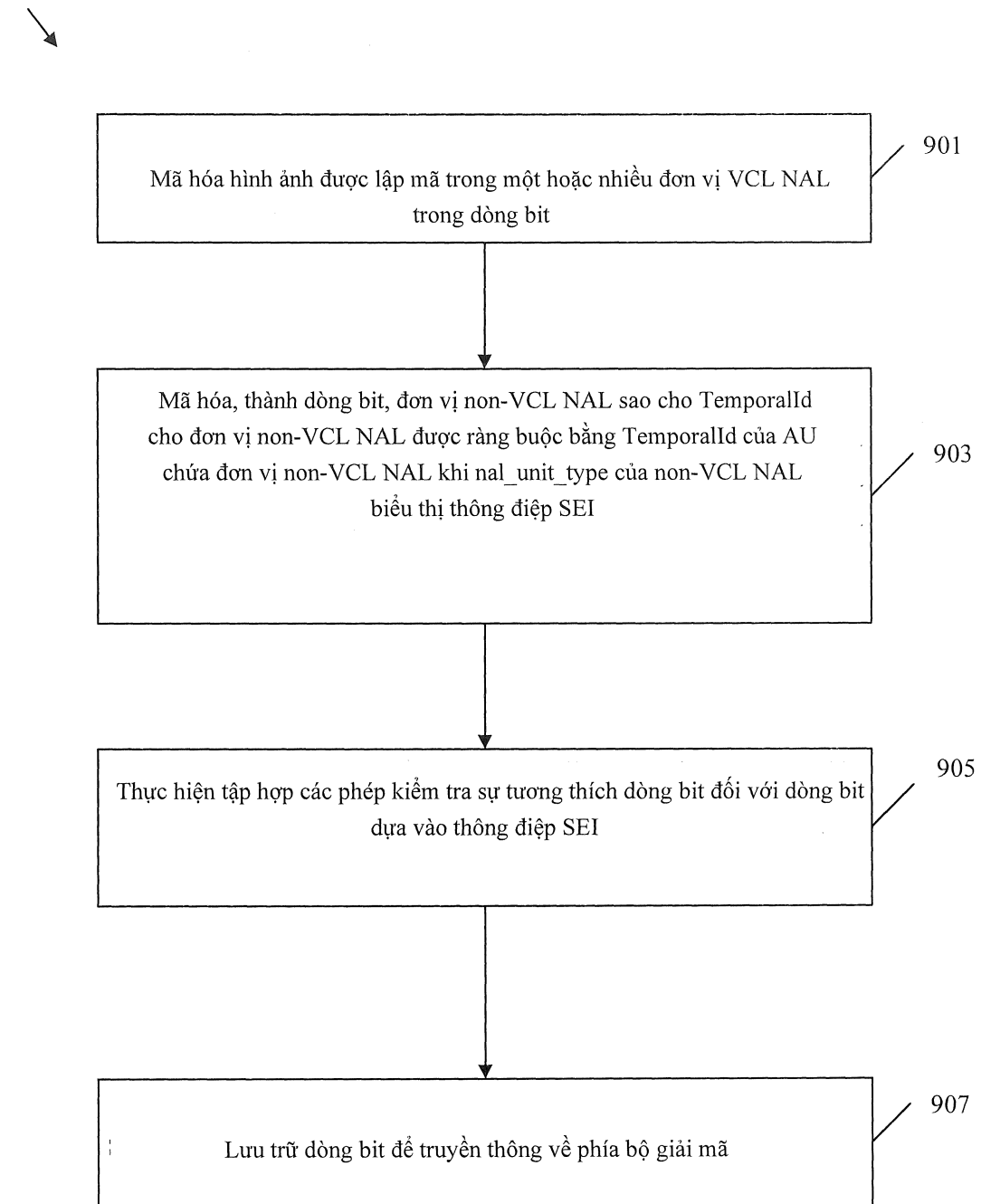


FIG. 9

10/11

1000

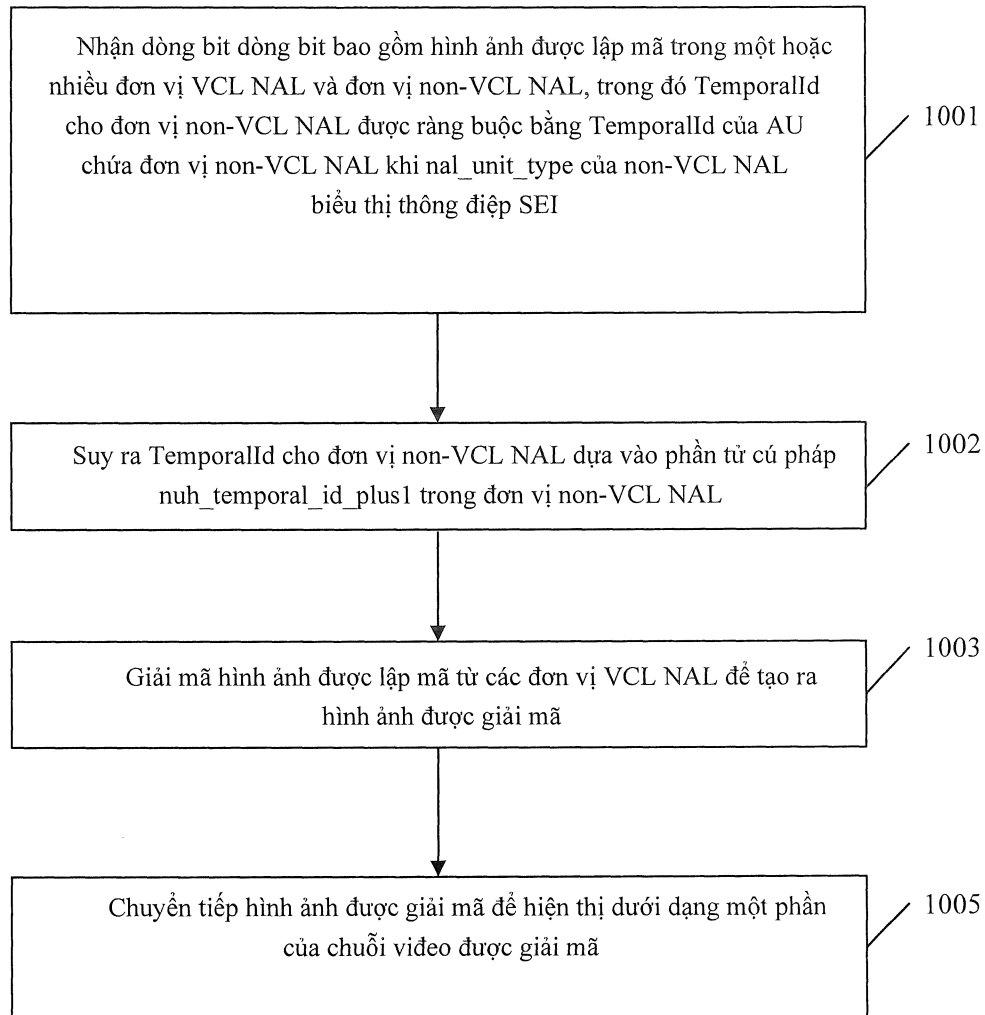


FIG. 10

11/11

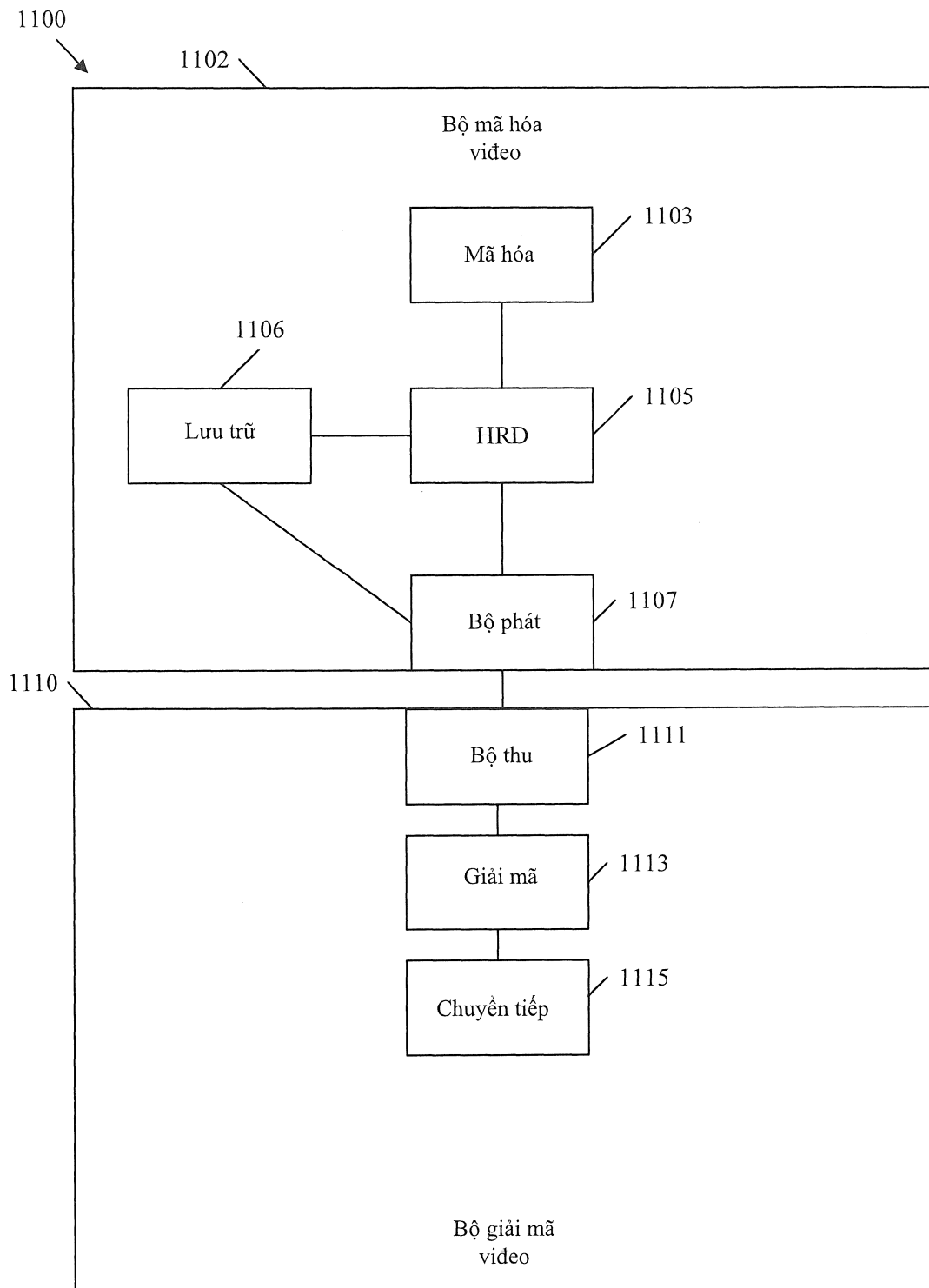


FIG. 11